



⑮ **BUNDESREPUBLIK  
DEUTSCHLAND**



**DEUTSCHES  
PATENT- UND  
MARKENAMT**

⑫ **Offenlegungsschrift**  
⑩ **DE 199 15 178 A 1**

⑤① Int. Cl. 7:  
**C 12 N 5/10**  
C 12 N 7/01  
A 61 K 48/00

②① Aktenzeichen: 199 15 178.4  
②② Anmeldetag: 3. 4. 1999  
④③ Offenlegungstag: 5. 10. 2000

**DE 199 15 178 A 1**

⑦① Anmelder:  
Johannes-Gutenberg-Universität Mainz, 55122  
Mainz, DE  
  
⑦④ Vertreter:  
Rudolph, U., Dipl.-Biol. Dr.rer.nat., Pat.-Anw., 69198  
Schriesheim

⑦② Erfinder:  
Bartenschlager, Ralf, Dr., 55239 Gau-Odernheim,  
DE  
  
⑤⑥ Entgegenhaltungen:  
J. gen. Virol. 78, S. 2397-2410, 1997;  
Virology 202, S. 606-614, 1994;

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤④ Hepatitis C Virus Zellkultursystem

⑤⑦ Das erfindungsgemäße Hepatitis C Virus (HCV) Zellkultursystem besteht aus humanen Hepatomazellen, die mit einem HCV-RNA-Konstrukt transfiziert sind, das die HCV-spezifischen RNA-Abschnitte 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3'NTR und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

**DE 199 15 178 A 1**

Die Erfindung betrifft ein Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, d. h. die mit HCV-spezifischem Genmaterial transfiziert sind.

Das Hepatitis C Virus (HCV) ist eine der Hauptursachen chronischer und sporadischer Lebererkrankungen weltweit. Die meisten HCV-Infektionen verlaufen ohne erkennbare klinische Symptome, allerdings werden 80–90% der Infizierten dauerhafte Virusträger und bei 50% dieser dauerhaften Virusträger kommt es zu einer chronischen Leberentzündung mit unterschiedlichen Schweregraden. Ca. 20% der chronisch Infizierten entwickeln im Laufe von 10 bis 20 Jahren eine Leberzirrhose, auf deren Basis sich ein primäres Leberzellkarzinom entwickeln kann. Die chronische Hepatitis C ist heute die Hauptindikation für eine Lebertransplantation. Eine Kausaltherapie gibt es bisher noch nicht. Die einzige derzeit verfügbare Therapie ist die hochdosierte Verabreichung von Interferon-Alpha oder eine Kombination aus Interferon-Alpha und dem Purin-Nukleosidanalogon Ribavirin. Allerdings sprechen nur ca. 60% aller Behandelten auf diese Therapie an und bei diesen kommt es in mehr als der Hälfte aller Fälle nach dem Absetzen der Behandlung zu einer erneuten Virämie.

Aufgrund der hohen Prävalenz, gerade auch in den Industrieländern, den schwerwiegenden Folgen chronischer Infektionen und dem Fehlen einer Kausaltherapie ist die Entwicklung einer HCV-spezifischen Chemotherapie ein wesentliches Ziel der pharmazeutischen Forschung und Entwicklung. Hauptproblem hierbei ist bisher das Fehlen eines geeigneten Zellkultursystems, das ein Studium der Virus-Replikation und der Pathogenese in eukaryontischen Zellen ermöglicht.

Aufgrund der geringen Virusmengen im Blut bzw. Gewebe, dem Fehlen geeigneter Zellkultursysteme oder Tiermodelle (bis heute ist der Schimpanse das einzige mögliche Versuchstier) sowie dem Fehlen effizienter Systeme zur Produktion virus-ähnlicher Partikel, konnte die molekulare Zusammensetzung des HCV-Partikels bis heute noch nicht eingehend untersucht bzw. aufgeklärt werden. Die derzeit vorliegenden Ergebnisse lassen sich wie folgt zusammenfassen: Das HCV ist ein umhülltes Plusstrang RNA Virus mit einem Partikeldurchmesser von 50–60 nm und einer mittleren Dichte von 1,03–1,1 g/ml. Es wurde erstmals 1989 molekular kloniert und charakterisiert (Choo et al., 1989: Science, 244, 359–362). Die HCV-RNA hat eine Länge von ca. 9.6 kb (= 9600 Nukleotide), eine positive Polarität und besitzt ein einziges offenes Leseraster (ORF = open reading frame), das ein lineares Polyprotein von ca. 3010 Aminosäuren kodiert (siehe Rice 1996, in Virology, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp. 931–960; Clarke 1997, J. Gen. Virol. 78, 2397; und Bartenschlager 1997, Intervirology 40, 378 und vgl. Fig. 1A). Bei der Virusreplikation wird das Polyprotein durch zelluläre und virale Proteasen in die reifen und funktionell aktiven Proteine gespalten.

Innerhalb des Polyproteins sind die Proteine wie folgt angeordnet (vom Amino- zum Carboxyterminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. Das Core-Protein ist die Hauptkomponente des Nukleokapsids. Die Glykoproteine E1 und E2 sind Transmembranproteine und Hauptkomponenten der Virushülle. Sie spielen wahrscheinlich bei der Anheftung des Virus an die Wirtszelle eine wesentliche Rolle. Diese drei Proteine Core, E1 und E2 bauen den Viruspartikel auf und werden deshalb als Strukturproteine bezeichnet. Die Funktion des Proteins p7 ist noch unklar. Das Protein NS2 ist wahrscheinlich die katalytische Domäne der NS2-3 Protease, die für die Prozessierung zwischen den Proteinen NS2 und NS3 verantwortlich ist. Das Protein NS3 hat zwei Funktionen, nämlich in der aminoterminalen Domäne eine Proteaseaktivität, die für die Polyproteinprozessierung essentiell ist, und in der carboxyterminalen Domäne eine NTPase/Helikase-Funktion, die wahrscheinlich bei der Replikation der viralen RNA eine Rolle spielt. Das Protein NS4A ist ein Kofaktor der NS3-Protease. Die Funktion des Proteins NS4B ist unbekannt.

Das offene Leseraster ist an seinem 5'Ende von einer ca. 340 Nukleotide langen nichttranslatierten Region (NTR = non-translated region) flankiert, die als interne Ribosomenansatzstelle (IRES = internal ribosome entry site) fungiert, und an seinem 3'Ende von einer ca. 230 Nukleotide langen NTR, die höchstwahrscheinlich für die Genomreplikation von Bedeutung ist. Eine solche 3'NTR ist Gegenstand der Patentanmeldung PCT/US 96/14033. Die Strukturproteine in dem aminoterminalen Viertel des Polyproteins werden von der Signalpeptidase der Wirtszelle gespalten. Die Nicht-Strukturproteine (NS) 2 bis (NS) 5B werden von zwei viralen Enzymen prozessiert, nämlich von der NS2-3 und der NS3/4A Proteinase. Die NS3/4A Proteinase wird für alle Spaltungen jenseits des Carboxyterminus von NS3 benötigt. Die Rolle von NS4B ist nicht bekannt. NS5A, ein hoch phosphoryliertes Protein, scheint für die Interferon Resistenz verschiedener HCV-Genotypen verantwortlich zu sein (vgl. Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187) und NS5B wurde als die RNA-abhängige RNA Polymerase identifiziert.

Anhand dieser Erkenntnisse wurden erste Diagnosesysteme entwickelt, die entweder auf dem Nachweis von HCV-spezifischen Antikörpern in Patientenserum oder auf dem Nachweis von HCV-spezifischer RNA mittels RT-PCR (= Reverse Transcription Polymerase Chain Reaction) beruhen, und die mittlerweile routine- und/oder vorschriftsmäßig bei allen Blutkonserven angewendet werden (müssen).

Seit der Erstbeschreibung des Genoms 1989 wurden mit Hilfe der PCR-Methode zahlreiche Teil- und Komplettssequenzen des HCV kloniert und charakterisiert. Ein Vergleich dieser Sequenzen zeigt eine hohe Variabilität des viralen Genoms, insbesondere im Bereich des NS5B-Gens, was letztendlich zu einer Einteilung in 6 Genotypen geführt hat, die selbst nochmals in Subtypen a, b, und c untergliedert sind. Die genomische Varianz ist nicht gleichmäßig über das Genom verteilt. So sind die 5'NTR und Teile der 3'NTR hoch konserviert, während bestimmte kodierende Sequenzen z. T. sehr stark variieren, vor allem die Hüllproteine E1 und E2.

Die klonierten und charakterisierten Teil- und Komplettssequenzen des HCV-Genoms wurden außerdem hinsichtlich geeigneter Angriffsziele für ein prospektives antivirales Therapeutikum untersucht. Dabei wurden drei virale Enzyme gefunden, die sich als solches Angriffsziel anbieten. Diese sind (1) der NS3/4A Proteasekomplex, (2) die NS3 Helikase und (3) die NS5B RNA-abhängige RNA Polymerase. Der NS3/4A Proteasekomplex und die NS3 Helikase konnten bereits kristallisiert und hinsichtlich ihrer dreidimensionalen Struktur aufgeklärt werden (Kim et al., 1996, Cell, 87, 343;

Ycm et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463; Cho et al., 1998, J. Biol. Chem., 273, 15045); bei der NS5B RNA-abhängigen RNA Polymerase ist dies bis heute noch nicht gelungen.

Obwohl mit diesen Enzymen bedeutsame Angriffsziele für eine Therapieentwicklung der chronischen HCV-Infektion definiert sind, und obwohl sowohl mit Hilfe von "rational drug design" als auch mit Hilfe von "high throughput screens" weltweit intensiv nach geeigneten Inhibitoren gesucht wird, leidet die Therapieentwicklung an einem großen Defizit, nämlich dem Fehlen von Zellkultursystemen oder einfachen Tiernodellen, die es erlauben, HCV-RNA oder HCV-Antigene direkt, zuverlässig und mit einfachen laborüblichen Methoden nachzuweisen. Das Fehlen solcher Zellkultursysteme ist auch der Hauptgrund dafür, daß das Verständnis der HCV-Replikation bis heute noch sehr lückenhaft und in weiten Teilen nur hypothetisch ist.

Obwohl nach Meinung der Fachwelt eine enge evolutionäre Beziehung zwischen HCV und den Flavi- und Pestiviren besteht und für diese autonom replizierende RNAs beschrieben sind, die in verschiedenen Zelllinien ohne weiteres zur Replikation gebracht werden können und dabei relativ hohe Ausbeuten zeigen (siehe Khromykh et al., 1997, J. Virol. 71, 1497; Behrens et al., 1998, J. Virol. 72, 2364; Moser et al., 1998, J. Virol. 72, 5318), waren ähnliche Versuche mit HCV bisher nicht erfolgreich.

Zwar ist aus verschiedenen Publikationen bekannt, daß Zelllinien oder primäre Zellkulturen mit HCV-haltigem, hochtitrigem Patientenserum infiziert werden können, (Lanford et al. 1994, Virology 202, 606; Shimizu et al. 1993, Proceedings of the National Academy of Sciences, USA, 90, 6037-6041; Mizutani et al. 1996, Journal of Virology, 70, 7219-7223; M. Ikeda et al. 1998, Virus Res. 56, 157; Fournier et al. 1998, J. Gen. Virol. 79, 2376 und darin zitierte Literaturstellen, Ito et al. 1996, Journal of General Virology, 77, 1043-1054), diese virusinfizierten Zelllinien oder Zellkulturen erlauben jedoch nicht den direkten Nachweis von HCV-RNA oder HCV-Antigenen. Die virale RNA in diesen Zellen ist weder in einem Northern-Blot (einem Standardverfahren zum quantitativen Nachweis von RNA) noch sind die viralen Proteine in einem Western-Blot oder mittels Immunpräzipitation detektierbar. Nur mit sehr aufwendigen und indirekten Methoden ist es überhaupt gelungen, eine HCV-Replikation nachzuweisen. Diese nachteiligen Umstände zeigen klar, daß die Replikation in diesen bekannten virusinfizierten Zelllinien oder Zellkulturen absolut unzureichend ist.

Desweiteren ist aus den Publikationen von Yoo et al. (1995, Journal of Virology, 69, 32-38) und von Dash et al., (1997, American Journal of Pathology, 151, 363-373) bekannt, daß Hepatomazelllinien mit synthetischer HCV-RNA, die mittels in vitro Transkription von kloniertem HCV-Genom gewonnen wurde, transfiziert werden können. In beiden Publikationen gingen die Autoren von dem Grundgedanken aus, daß das virale HCV-Genom eine Plusstrang-RNA ist, die nach dem Einschleusen in die Zelle direkt als mRNA fungiert, an die sich Ribosomen anheften und im Zuge von Translationsprozessen Virusproteine bilden, aus denen sich letztendlich neue HCV-Partikel bilden (können). Diese Virusreplikation, d. h. diese neu gebildeten HCV-Viren bzw. deren RNA wurde mittels RT-PCR nachgewiesen. Die publizierten Ergebnisse der durchgeführten RT-PCR sprechen jedoch dafür, daß die Effizienz der HCV-Replikation in den beschriebenen HCV-transfizierten Hepatomazellen nur sehr gering ist und jedenfalls nicht ausreicht, um Schwankungen in der Replikationsrate nach gezielter Einwirkung mit prospektiven antiviralen Therapeutika auch nur qualitativ, geschweige denn quantitativ zu messen.

Aufgabe der vorliegenden Erfindung ist die Bereitstellung eines HCV-Zellkultursystems, bei dem die virale RNA in den transfizierten Zellen autonom und mit so hoher Effizienz repliziert, daß Schwankungen in der Replikationsrate nach gezielter Einwirkung mit virus- und insbesondere HCV-spezifischen prospektiven antiviralen Therapeutika qualitativ und quantitativ und mit Hilfe gängiger, laborüblicher Meßverfahren gemessen werden können.

Eine Lösung dieser Aufgabe besteht in der Bereitstellung eines Zellkultursystems der eingangs genannten Art, bei dem die eukaryontischen Zellen humane Hepatomazellen sind, die vorzugsweise von einer handelsüblichen Hepatomazelllinie abstammen, aber auch aus einer entsprechenden Primärzellkultur gewonnen sein können, und bei dem das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3'NTR, vorzugsweise in der genannten Reihenfolge, und zudem wenigstens ein selektierbares Markergen (Selektionsgen) umfaßt.

"NTR" steht hier und im folgenden für "nicht translatierte Region" und ist dem einschlägigen Fachmann als Begriff bzw. Abkürzung bekannt und geläufig.

Der Begriff "HCV-RNA-Konstrukt" umfaßt hier und im folgenden sowohl Konstrukte, die das komplette HCV-Genom enthalten, als auch solche, die lediglich einen Teil davon, d. h. ein HCV-Subgenom enthalten.

Eine bevorzugte Variante des erfindungsgemäßen Zellkultursystems, die sich in der Praxis sehr gut bewährt hat, trägt ist unter der Nummer DSM ACC2394 (Laborbezeichnung HuB1 9-13) bei der DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH in Braunschweig, Deutschland, hinterlegt.

Mit dem erfindungsgemäßen Zellkultursystem wird erstmals ein in-vitro-System bereit gestellt, in dem HCV-RNA intrazellulär, autonom und in ausreichend großen Mengen repliziert und exprimiert wird, so daß eine quantitative Bestimmung sowohl der HCV-RNA-Mengen als auch der HCV-spezifischen Proteine mit konventionellen und zuverlässig genauen biochemischen Meßmethoden durchgeführt werden kann. Das heißt: es steht erstmals ein annähernd authentisches zellgestütztes ("cell-based") HCV-Replikationssystem zur Verfügung, das für die Entwicklung und Erprobung von antiviralen Pharmazeutika dringend benötigt wird. Dieses Testsystem bietet nun die Möglichkeit, potentielle Angriffsziele für eine wirksame HCV-spezifische Therapie zu identifizieren und HCV-spezifische Chemotherapeutika zu entwickeln und zu evaluieren.

Die Erfindung basiert auf der überraschenden Erkenntnis, daß eine effiziente Replikation der HCV-RNA nur dann in Zellen stattfindet, wenn diese mit einem HCV-RNA-Konstrukt transfiziert wurden, das mindestens die 5' und die 3' nicht-translatierten Regionen (NTR) und die Nichtstrukturproteine (NS) 3 bis 5B umfaßt und zusätzlich ein selektierbares Markergen (Selektionsgen) aufweist. Offensichtlich sind die Strukturgene für den Ablauf der Replikation ohne wesentliche Bedeutung, während andererseits eine effiziente Replikation der HCV-RNA anscheinend nur dann stattfindet, wenn die transfizierten Zellen einem permanenten Selektionsdruck unterzogen werden, der durch das mit der HCV-RNA verbundene selektierbare Markergen (Selektionsgen) vermittelt wird. Das Markergen (Selektionsgen) scheint somit einer-

seits die Selektion derjenigen Zellen zu provozieren, in denen die HCV-RNA produktiv repliziert, und andererseits scheint es die Effizienz der RNA-Replikation wesentlich zu steigern.

Gegenstand der Erfindung ist auch ein zellfreies HCV-RNA-Konstrukt, das sich dadurch auszeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3'NTR, vorzugsweise in der genannten Reihenfolge, und zudem ein selektierbares Markergen (Selektionsgen) umfaßt.

Der Begriff 5'NTR bzw. NS3 bzw. NS4A bzw. NS4B bzw. NS5A bzw. NS5B bzw. 3'NTR umfaßt im vorliegenden Zusammenhang jede Nukleotidsequenz, die im Stand der Technik als Nukleotidsequenz für den jeweils betreffenden funktionellen Abschnitt des HCV-Genoms beschrieben ist.

Die Bereitstellung eines solchen HCV-RNA-Konstrukts ermöglicht erstmals eine detaillierte Analyse der HCV-Replikation, -Pathogenese und -Evolution in Zellkulturen. Die HCV-spezifische virale RNA kann – als vollständiges Genom oder als Subgenom – gezielt in beliebigen Mengen erzeugt werden, und es besteht die Möglichkeit, das RNA-Konstrukt zu manipulieren und damit die HCV-Funktionen auf genetischer Ebene zu untersuchen und aufzuklären.

Da alle zur Zeit als Hauptangriffsziel für eine Therapie untersuchten HCV-Enzyme, nämlich die NS3/4A Protease, die NS3 Helikase und die NS5B Polymerase, in dem erfindungsgemäßen HCV-RNA-Konstrukt enthalten sind, kann es für alle entsprechenden Untersuchungen benutzt werden.

Eine Ausführungsform des HCV-RNA-Konstrukts, die sich in der praktischen Anwendung sehr gut bewährt hat, zeichnet sich dadurch aus, daß sie die Nukleotidsequenz gemäß Sequenzprotokoll SEQ ID NO: 1 umfaßt.

Weitere Ausführungsvarianten mit vergleichbar guten Eigenschaften für den Einsatz in der Praxis sind dadurch gekennzeichnet, daß sie eine Nukleotidsequenz entweder gemäß Sequenzprotokoll SEQ ID NO: 2 oder gemäß Sequenzprotokoll SEQ ID NO: 3 oder gemäß Sequenzprotokoll SEQ ID NO: 4 oder gemäß Sequenzprotokoll SEQ ID NO: 5 umfassen.

Es besteht die Möglichkeit, das erfindungsgemäße HCV-Subgenom-Konstrukt mit einer 3'NTR zu versehen, die eine im Stand der Technik hierfür bisher unbekannte Nukleotidsequenz aufweist, nämlich eine Nukleotidsequenz, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

(a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTTTA  
GCTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCC  
AGACTGGAGT

(b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT  
TGCCAGACT GGAGC

(c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTAATCTTT TTTTTTTTTCT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT  
TGCCAGACT GCAGC

(d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTTAGTC TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACTC  
TGTTGCCCAG ACTGGAGT

(e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACTC  
TGTTGCCCAG ACTGGAGT

(f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TTGAGAGAGA GAGTCTCACT  
CTGTTGCCCA GACTGGAGT

- (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA  
GTCTCACTCT GTTGCCCACTGAGT
- (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTTTAAT CTTTTTTTTT TTTTTCCTTT TTTTGAGAGA  
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T
- (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT  
TTTTTTAATC TTTTTTTTTT TTTTCTTTT TTTTGAGAG  
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT

Das in den erfindungsgemäßen HCV-RNA-Konstrukten enthaltene selektierbare Markergen (Selektionsgen) ist vorzugsweise ein Resistenzgen, insbesondere eine Antibiotikumresistenzgen.

Das hat den Vorteil, daß die mit diesem Konstrukt transfizierten Zellen leicht von den nicht transfizierten Zellen selektiert werden können, indem dem Zellkulturmedium z. B. im Fall eines Antibiotikumresistenzgens das betreffende Antibiotikum zugegeben wird. Unter "Antibiotikum" wird im vorliegenden Zusammenhang jede Substanz verstanden, die die nicht-transfizierten Wirtszellen oder die Zellen, in denen die HCV-RNA nur mit geringer Effizienz repliziert, am Leben oder Wachstum hindert, insbesondere Zellgifte wie z. B. Puromycin, Hygromycin, Zeocin, Bleomycin oder Blastidin.

Ein bevorzugtes selektierbares Markergen (Selektionsgen) bzw. Resistenzgen, das sich in der Praxis sehr gut bewährt hat, ist das Neomycinphosphotransferasegen.

Eine Alternative zu den Antibiotikumresistenzgenen ist z. B. das Thymidin-Kinase-Gen, mit dem eine HAT-Selektion durchgeführt werden kann.

Die Position des selektierbaren Markergens (Selektionsgens), bzw. des bevorzugten Resistenzgens bzw. des besonders bevorzugten Antibiotikumresistenzgens in dem HCV-RNA-Konstrukt liegt vorzugsweise hinter der HCV 5'NTR, d. h. strangabwärts der 5'NTR bzw. strangaufwärts des HCV-Leserasters. Denkbar ist aber auch eine Insertion im Bereich der 3'NTR oder an anderer Stelle des HCV-Genoms oder -Subgenoms, z. B. innerhalb des Polyproteins.

Bei einer alternativen Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts ist das selektierbare Markergen (Selektionsgen), insbesondere ein Antibiotikumresistenzgen, über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA bzw. der HCV-Genom- oder -Subgenomsequenz verbunden.

Damit geht der Vorteil einher, daß nach erfolgter Selektion derjenigen Zellen, in denen die HCV-RNA produktiv repliziert, in den daraus gewonnenen Zellklonen das Resistenzgen durch ribozymvermittelte Spaltung von der HCV-Subgenomsequenz abgetrennt werden kann, nämlich durch Aktivierung des einklonierten Ribozyms oder, im Fall eines Konstrukts mit einer Erkennungsstelle für ein Ribozym, durch Einschleusen des Ribozyms in die Zellen (z. B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). Auf diese Weise wird ein authentisches HCV-Genom-Konstrukt ohne Resistenzgen erhalten, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist.

Eine weitere bevorzugte Ausführungsform des erfindungsgemäßen HCV-RNA-Konstrukts zeichnet sich dadurch aus, daß das Konstrukt wenigstens ein integriertes Reportergen aufweist.

Unter Reportergen wird im folgenden jedes Gen verstanden, dessen Anwesenheit sich nach Überführung in einen Zielorganismus leicht und im allgemeinen mit einfachen biochemischen oder auch histochemischen Methoden nachweisen läßt, d. h. das für ein Protein kodiert, welches auch in geringen Mengen einfach und zuverlässig mit den laborüblichen Meßmethoden nachgewiesen und quantifiziert werden kann.

Diese Variante des HCV-RNA-Konstrukts hat den Vorteil, daß der Umfang der Replikation dieses Konstrukts anhand des Reportergenprodukts einfach und schnell mit laborüblichen Methoden gemessen werden kann.

Das Reportergen ist vorzugsweise ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), dem GFP-Gen (green-fluorescence-protein-Gen), dem GUS-Gen (Glukuronidasegen) oder dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen). Diese Reportergene bzw. deren Produkte, nämlich die entsprechenden Reporterproteine, können z. B. mittels Fluoreszenz, Chemilumineszenz, colorimetrisch oder mit Hilfe immunologischer Methoden (z. B. ELISA) bestimmt werden.

Als Reportergen kommt aber auch ein Surrogatmarkergen in Betracht. Darunter sind in diesem Zusammenhang solche Gene zu verstehen, die für zelluläre Proteine, Nukleinsäuren oder – allgemein – für solche Funktionen kodieren, die einer von der Virusreplikation abhängigen Variation unterliegen, und die infolgedessen in denjenigen Zellen, in denen sich das HCV bzw. das HCV-RNA-Konstrukt vermehrt, entweder reprimiert oder aktiviert werden. Das heißt: die Reduktion bzw. Aktivierung dieser Funktion ist ein Ersatzmarker für die Virusreplikation bzw. die Replikation des HCV-RNA-Konstrukts.

Die Positionen von Reportergen und selektierbarem Markergen (Selektionsgen) können so gewählt sein, daß ein aus den beiden Genprodukten gebildetes Fusionsprotein exprimiert wird. Hierbei besteht die vorteilhafte Möglichkeit, daß diese beiden Gene so in dem HCV-RNA-Konstrukt angeordnet sind, daß ihre beiden exprimierten Proteine zunächst über eine Schnittstelle für eine Protease (z. B. Ubiquitin) oder über ein selbstspaltendes Peptid (z. B. das 2A-Protein der Picornaviren) fusioniert sind und erst später proteolytisch wieder getrennt werden.

Ebensogut können diese beiden Positionen aber auch derart getrennt voneinander liegen, daß beide Genprodukte separat exprimiert werden. (z. B. in der Reihenfolge: Marker- bzw. Resistenzgen – interne Ribosomenbindungsstelle – Re-

portergen).

Im Fall des Reportergens hat sich eine Ausführungsvariante besonders bewährt, bei der das Reportergen in das offene Leseraster des HCV-Genoms oder -Subgenoms inkloniert ist, und zwar derart, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführt wird.

- 5 Das erfindungsgemäße Zellkultursystem in allen seinen Variationen kann für vielfältige Zwecke eingesetzt werden. Diese umfassen:

- Das Auffinden antiviral wirksamer Substanzen. Dies können beispielsweise sein: organische Verbindungen, die unmittelbar oder mittelbar in die Virusvermehrung eingreifen (z. B. Inhibitoren der viralen Proteasen, der NS3-Helikase, der NS5B RNA-abhängigen RNA Polymerase), antisense Oligonukleotide, die an eine beliebige Zielsequenz innerhalb des HCV-RNA-Konstrukts (z. B. die 5'NTR) hybridisieren und unmittelbar oder mittelbar zu einer Beeinflussung der Virusvermehrung führen z. B. auf Grund einer Reduktion der Translation des HCV-Polypoteins oder Ribozyme, die eine beliebige HCV-RNA-Sequenz spalten und damit die Virusreplikation beeinträchtigen.
- Die Evaluierung jeglicher Art antiviral wirksamer Substanzen in Zellkultur. Solche Substanzen können beispielsweise mittels "rational drug design" oder "high-throughput screening" am isolierten gereinigten Enzym gefunden werden. Unter Evaluierung sind vor allem die Bestimmung der inhibitorischen Eigenschaften der entsprechenden Substanz sowie deren Wirkungsmechanismus zu verstehen.
- Die Identifikation neuer Angriffsziele, viralen oder zellulären Ursprungs, für eine HCV-spezifische antivirale Therapie. Ist beispielsweise ein zelluläres Protein essentiell für die Virusreplikation, kann mittels Hemmung dieses zellulären Proteins die Virusreplikation ebenfalls beeinflusst werden. Das Auffinden solcher auxiliären Faktoren ist mit dem erfindungsgemäßen System ebenfalls möglich.
- Der Einsatz für die Resistenzbestimmung. Es ist anzunehmen, daß auf Grund der hohen Mutationsrate des HCV-Genoms Therapieresistenzen auftreten. Solche Resistenzen, die gerade bei der klinischen Zulassung einer Substanz von großer Bedeutung sind, lassen sich mit dem erfindungsgemäßen Zellkultursystem ermitteln. Zelllinien in denen sich das HCV-RNA-Konstrukt bzw. das HCV-Genom oder -Subgenom repliziert, werden mit steigenden Konzentrationen der entsprechenden Substanz inkubiert und die Replikation der viralen RNA wird entweder anhand eines eingebrachten Reporters oder durch qualitative oder quantitative Bestimmung der viralen Nukleinsäuren oder Proteine bestimmt. Resistenz ist dann gegeben, wenn bei normaler Wirkstoffkonzentration keine Hemmung der Replikation zu beobachten ist. Durch Reklonierung der HCV-RNA (z. B. mittels RT-PCR) und Sequenzanalyse können die für Therapieresistenz verantwortlichen Nukleotid- bzw. Aminosäureaustausche ermittelt werden. Durch Einklonieren der/des entsprechenden Austausches in das Ursprungs-konstrukt kann deren Kausalität für die Therapieresistenz bewiesen werden.
- Die Produktion authentischer Virusproteine (Antigene) für die Entwicklung und/oder Evaluierung von Diagnostika. Das erfindungsgemäße Zellkultursystem erlaubt auch die Expression von HCV-Antigenen in Zellkulturen. Diese Antigene können prinzipiell auch für den Aufbau diagnostischer Nachweisverfahren eingesetzt werden.

Das erfindungsgemäße HCV-RNA-Konstrukt für sich genommen kann in allen seinen Variationen ebenfalls für vielfältige Zwecke eingesetzt werden. Dazu gehören vor allem:

- Die Konstruktion attenuierter Hepatitis C Viren bzw. HCV-ähnlicher Partikel und deren Produktion in Zellkulturen:  
Durch zufällige oder gezielt hervorgerufene Mutationen, beispielsweise Punktmutationen, Deletionen oder Insertionen, können attenuierte HCV- oder HCV-ähnliche Partikel erzeugt werden, d. h. Viren bzw. virusähnliche Partikel mit voller Replikationskompetenz aber verringerter bzw. fehlender Pathogenität. Solche attenuierte HCV- oder HCV-ähnliche Partikel sind insbesondere als Impfstoff einsetzbar.
- Die Konstruktion von HCV-RNA-Konstrukten mit integrierten Fremdgenen, beispielsweise zur Verwendung als leberzellspezifische Gentherapien in der Genterapie. Auf Grund des ausgeprägten Leberzelltropismus des HCV und der Möglichkeit, Teile des Genoms durch heterologe Sequenzen zu ersetzen, lassen sich HCV-RNA-Konstrukte herstellen, bei denen beispielsweise die Strukturproteine durch ein therapeutisch wirksames Gen ersetzt werden. Das so erhaltene HCV-RNA-Konstrukt wird in Zellen eingeschleust, vorzugsweise mittels Transfektion, die die fehlenden HCV-Funktionen, beispielsweise die Strukturproteine, konstitutiv oder induzierbar exprimieren. Durch diese dem Fachmann unter dem Begriff der "Transkomplementation" bekannte Technik lassen sich Viruspartikel erzeugen, in die das HCV-RNA-Konstrukt eingebaut wird. Die so erhaltenen Partikel können für die Infektion vorzugsweise von Leberzellen verwendet werden. In diesen wird das therapeutisch wirksame Fremdgen zur Expression gebracht und entfaltet damit seine therapeutische Wirkung.
- Das Auffinden permissiver Zellen, d. h. Zellen, in denen eine produktive Virusvermehrung erfolgt. Zu diesem Zweck wird entweder eines der vorgenannten HCV-RNA-Genomkonstrukte verwendet, das zur Bildung kompletter infektiöser Viren befähigt ist, oder es wird eines der vorgenannten HCV-Subgenom-Konstrukte eingesetzt, das allerdings zunächst gemäß vorgenanntem Beispiel in eine Zelllinie transfiziert wird, die die fehlenden Funktionen konstitutiv oder induzierbar exprimiert. In all diesen Fällen entstehen Viruspartikel, die zusätzlich zur HCV-Sequenz ein Resistenz- und/oder Reportergen tragen. Zum Auffinden von Zellen, in denen das HCV replizieren kann, werden diese Zellen mit den so hergestellten Viren infiziert und einer Antibiotikumsselektion unterzogen oder, in Abhängigkeit vom HCV-RNA-Konstrukt, mittels Nachweis der Expression des Reportergens untersucht. Da eine Antibiotikumsresistenz bzw. eine Expression des Reportergens nur dann nachweisbar ist, wenn das HCV-RNA-Konstrukt repliziert, müssen die so gefundenen Zellen permissiv sein. Auf diese Weise lassen sich nahezu beliebige Zelllinien oder primäre Zellkulturen hinsichtlich der Permissivität testen und auffinden.

Die Erfindung wird im folgenden anhand von Ausführungsbeispielen und dazugehörigen Figuren näher erläutert. Es

zeigen

**Fig. 1A** die Struktur eines erfindungsgemäßen HCV-RNA-Konstrukts Ganz oben ist eine schematische Darstellung der Struktur des kompletten parentalen HCV-Genoms gegeben mit den Positionen der Gene für die Spaltungsprodukte core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A und NS5B innerhalb des Polyproteins, und den 5' und 3' nichttranslatierten Regionen (5'NTR und 3'NTR) – als Horizontalbalken dargestellt –, und mit den beiden für die Erzeugung der Subgenom-Konstrukte ausgewählten Positionen, nämlich der Position der "GDD-katalytischen Domäne" der NS5B RNA Polymerase (GDD) und der Position der 3'Grenze der HCV-IRES (Nukleotidpositionen 1 bis 377 bzw. 1 bis 389) – oberhalb des Genomschemas eingezeichnet –. Die Zahlen unterhalb des Genomschemas bezeichnen die entsprechenden Nukleotidpositionen.

Darunter sind schematische Darstellungen der Strukturen zweier erfindungsgemäßer, modifizierter HCV-RNA-Konstrukte (Subgenom) gezeigt, bestehend aus der 5'HCV-IRES, dem Neomycinphosphotransferasegen (Neo<sup>R</sup>), der EMCV-IRES (E-I) und den HCV Sequenzen von NS2 bzw. NS3 bis zum authentischen 3'Ende. Die Position der das NS5B Polymerase GDD-Motiv umfassenden 10 Aminosäuren-Deletion ist jeweils mit einem Dreieck ( $\Delta$ ) markiert;

**Fig. 1B** das Ergebnis einer denaturierenden Formaldehyd-Agarose-Gelelektrophorese zum Nachweis von replizierter Plusstrang-RNA in transfizierten subpassagierten Huh-7 Zellklonen.

Die Positionen der HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker (M) sind links von Spur 1 angegeben;

**Fig. 1C** das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum Nachweis der Abwesenheit von integrierter Replikon-DNA in den meisten der selektierten Zellklone.

Spuren 1 und 2 zeigen die Positivkontrollen, Spur 13 die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle;

**Fig. 2A** das Ergebnis eines PCR-Tests mit nachfolgendem Southern-Blot zum sensitiven Ausschluß integrierter Replikon-DNA (Plasmid Moleküle I<sub>377</sub>/NS3-3'/wt) in einem HCV-RNA-Konstrukt-haltigen Zellklon (9-13).

Die Spuren 7 bis 11 repräsentieren das Ergebnis einer Titration von DNA-Molekülen des Konstrukts I<sub>377</sub>/NS3-3'/wt ohne Zugabe von Gesamt-DNA des Zellklons 9-13 und die Spuren 2-6 repräsentieren die gleichen Plasmidmoleküle mit Zugabe von jeweils 1 µg 9-13 DNA vor der PCR (zwecks Ausschluß eines Inhibitors der PCR in der DNA-Präparation). Spur 13 repräsentiert die Negativ-Kontrolle (PCR ohne DNA-Sonde). Spur 1 zeigt das Ergebnis, das mit einem µg Gesamt-DNA des Zellklons 9-13 erhalten wurde;

**Fig. 2B** das Ergebnis eines Northern-Blot-Tests zur Quantifizierung von HCV Plus- und Minusstrang RNA.

Die Pfeile markieren die Positionen von Replikon-RNA.

Die "plus"- und "minus"-Abgaben bezeichnen die positive (plus) bzw. negative (minus) Polarität der RNA-Kontrollen, die auf das Gel aufgetragen wurden. "Minusstrang" und "Plusstrang" bezeichnen die Spezifität der radioaktiven RNA-Sonden;

**Fig. 2C** Ergebnis einer Formaldehyd-Agarose-Gelelektrophorese nach radioaktiver Markierung der intrazellulär replizierten HCV-RNA zum Nachweis der Resistenz der HCV-RNA-Replikation gegen Dactinomycin;

**Fig. 3A** Nachweis von HCV-spezifischen Antigenen in den selektierten Zellklonen mittels Immunopräzipitation nach metabolischer Radioaktivmarkierung.

Die Spuren 7-9 repräsentieren authentische Größenmarker (die nach transienter Expression eines HCV-RNA-Konstrukts in Huh-7-Zellen erhalten wurden); identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben;

**Fig. 3B** Ergebnisse eines Immunfluoreszenztests zum Nachweis der subzellulären Lokalisation von HCV Antigenen;

**Fig. 4** schematische Darstellung der Struktur eines erfindungsgemäßen selektierbaren HCV-RNA-Konstrukts (komplettes Genom) bestehend aus der 5'HCV-IRES, dem Neomycinphosphotransferasegen (Neo<sup>R</sup>), einem heterologen IRES-Element, z. B. des Encephalomyocarditisvirus (E-I), dem vollständigen HCV-Leseraster und der authentischen 3'NTR;

**Fig. 5** schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit inseriertem Antibiotikumresistenzgen (A) innerhalb der für das Polyprotein kodierenden Nukleotidsequenz (monocistronische RNA), und (B) innerhalb der 3'NTR (bicistronische RNA);

**Fig. 6** schematische Darstellung der Struktur von HCV-RNA-Konstrukten mit inseriertem Reportergen (A) als Teil eines HCV-Replikons von NS3 bis NS5B; – das Reportergen wird letztendlich durch virale oder durch zelluläre Proteasen aus dem Polyprotein gespalten und das selektierbare Markergen (Selektionsgen) bzw. das Resistenzgen durch Kontransfektion in die Zellen einschleust, (B) als Teil eines Fusionsgens aus Resistenz- und Reportergen (z. B. für die Neomycinphosphotransferase und green fluorescent Protein) (C) als Teil eines Replikons aus Resistenz- und Reportergen (z. B. für die Neomycinphosphotransferase und das green fluorescent Protein), die über eine Nukleotidsequenz verbunden sind, welche für eine Aminosäuresequenz kodiert (schraffierter Bereich), die von einer Protease gespalten werden kann oder die über eine selbstspaltende (autokatalytische) Aktivität verfügt, ( $\Delta$ ) als unabhängiges Gen (hier green fluorescent protein), das von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert wird; – das Resistenzgen (hier: Neomycinphosphotransferase-Gen) wird davon unabhängig ebenfalls von einer eigenen internen Ribosomenbindungsstelle (IRES) aus exprimiert (polycistronisches Konstrukt);

**Fig. 7** schematische Darstellung der Struktur eines HCV-RNA-Konstrukts bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist.

Die dicken Linien stellen die HCV 5'- und 3'NTRs dar, E-I ist eine heterologe interne Ribosomenbindungsstelle, die für die Expression des Resistenzgens notwendig ist, und das graue Quadrat stellt das Ribozym bzw. eine Erkennungsstelle für ein Ribozym dar;

**Fig. 8** schematische Darstellung der Struktur eines HCV-RNA-Konstrukts mit Resistenzgen und integriertem Fremdgen.

## Herstellung von HCV-RNA-Konstrukten

## 5 (A) Synthese und Klonierung eines vollständigen HCV-Konsensusgenoms mittels RT-PCR

Aus der Leber eines chronisch infizierten Patienten wurde das HCV-Genom, d. h. die HCV-RNA wie nachfolgend beschrieben isoliert:

Aus ca. 100 mg Leber wurde die komplette RNA gemäß dem Verfahren von Chomczynski und Sacchi (1987, Anal. Biochem. 162, 156) isoliert. Mit 1 µg dieser isolierten RNA wurde eine reverse Transkription mit den Primern A6103. (GCTATCAGCCGGTTCATCCACTGC) oder A9413 (CAGGATGGCCTATTGG CCTGGAG) und dem "expand reverse transcriptase"-System (Boehringer Mannheim, Deutschland) nach den Vorschriften des Herstellers durchgeführt. Mit den Produkten dieser reversen Transkription (RT) wurde eine Polymerase-Kettenreaktion (PCR = polymerase chain reaction) durchgeführt, und zwar unter Verwendung des "expand long template"-Systems (Boehringer Mannheim, 15 Deutschland), wobei der Puffer mit 2% Dimethylsulfoxid-Gehalt eingesetzt wurde. Nach einer Stunde bei 42°C wurde 1/8 dieses Reaktionsansatzes in einem ersten PCR-Durchgang mit den Primern A6103 und S59 (TGTCTTCACGCA-GAAAGCGTCTAG) oder A9413 und 54542 (GATGAGCT CGCCGCGAAGCTGTCC) eingesetzt. Nach 40 Zyklen wurde 1/10 dieses Reaktionsansatzes in einem zweiten PCR-Durchgang mit den Primern S59 und A4919 (AGCA-CAGCCCGCGTCATAGCACTCG) oder 54542 und A9386 (TTAGTCTCCCG TTCATCGGTTGG) eingesetzt. Nach 20 30 Zyklen wurden die PCR-Produkte mittels präparativer Agarose-Gel-Elektrophorese gereinigt und die dabei eluierten Fragmente wurden in den Vektor pCR2.1 (Invitrogen) oder pBSK II (Stratagene) ligiert. Vier Klone von jedem Fragment wurden analysiert und sequenziert, und es wurde eine Konsensus-Sequenz ermittelt. Zu diesem Zweck wurden die DNA-Sequenzen miteinander verglichen. Die Positionen, an denen sich die Sequenz eines der Fragmente von den übrigen unterschied, wurde als unerwünschte Mutation betrachtet. Im Fall von Mehrdeutigkeiten der Sequenz wurden kürzere sich überlappende PCR-Fragmente der betreffenden Region amplifiziert und mehrere Klone sequenziert. Auf diese Weise 25 konnten zahlreiche potentielle Mutationen in jedem Fragment identifiziert und somit eine isolat-spezifische Konsensussequenz etabliert werden. Diese etablierte Konsensussequenz bzw. dieses Genom gehört zum weltweit verbreiteten Genotyp 1b. Die nicht translatierte Region am 3'-Ende (= 3'NTR) wurde mittels konventioneller PCR erhalten, wobei ein Antisense-Primer eingesetzt wurde, der die letzten 24 Nukleotide des im Stand der Technik bekannten "X-tails" (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033) abdeckt. Die authentische nicht translatierte Region am 5'-Ende (= 5'NTR) strangabwärts vom T7 Promotor wurde mittels PCR erzeugt, wobei zum einen ein Oligonukleotid verwendet wurde, das einem verkürzten T7 Promotor (TAA TAC GAC TCA CTA TAG) und den ersten 88 Nukleotiden von HCV entspricht, und zum anderen eines der vorgenannten Plasmide eingesetzt wurde, das eines der 5'Fragmente des Genoms trägt. Aus den subgenomischen Fragmenten mit der geringsten Anzahl an Nicht-Konsensus-Austauschen wurde ein komplettes HCV-Konsensusgenom zusammengesetzt und in einen modifizierten pBR322-Vektor inseriert. Abweichungen von der Konsensussequenz wurden mittels ortsgerichteter Mutagenese ("site-directed mutagenesis") beseitigt. Um "run-off"-Transkripte mit einem authentischen 3'Ende herzustellen, wurde die 3'NTR der Isolate (mit dem Ende TGT) zu AGT modifiziert (gemäß der Sequenz vom Genotyp 3 = Klon "WS" nach Kolykhalov et al., 1996, J. Virol. 70, 3363) und außerdem wurde ein zusätzlicher Nukleotidaustausch an Position 9562 vorgenommen, um die A:T Basenpaarung in der Haarnadelstruktur am 3'Ende der 3'NTR (Kolykhalov et al. *ibid.*) beizubehalten. Um eine interne Restriktionsstelle für das Enzym Scal zu beseitigen, wurde ferner ein sog. stiller ("silent") Nukleotidaustausch vorgenommen. Nach dem Zusammenfügen des Vollängen-Genoms mit passenden 5'- und 3'NTRen wurde die komplette HCV-Sequenz überprüft. Dabei wurde kein ungewünschter Nukleotidaustausch gefunden.

Das auf diese Weise hergestellte HCV-Genom sollte per Definition hepatotrop sein.

45

## (B) Synthese selektierbarer HCV-Subgenom-Konstrukte

Unter Verwendung des unter (A) beschriebenen Konsensusgenoms wurden HCV-Subgenom-Konstrukte hergestellt, die das Antibiotikumresistenzgen Neomycin-Phosphotransferase (NPT) und zwei Sequenzen von internen Ribosomenbindungsstellen (IRES) enthalten. Die hierfür angewendeten biochemischen Verfahrenstechniken sind dem Fachmann bekannt und geläufig (siehe: Sambrook, J., E. F. Fritsch, T. Maniatis, 1989, Molecularcloning: a laboratory manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N. Y.; Ausubel et al. (eds.), 1994, Current Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons Inc., New York). Das Antibiotikumresistenzgen wurde unmittelbar hinter der 5'NTR inseriert, wodurch eine bicistronische RNA erhalten wurde (siehe Fig. 1A). Ebenso gut kann das Antibiotikumresistenzgen aber auch an anderer Stelle des HCV-Subgenom-Konstrukts inseriert werden, beispielsweise innerhalb der für das Polyprotein kodierenden Nukleotidsequenz, wodurch eine monocistronische RNA erhalten wird (siehe Fig. 5A) oder in die 3'NTR (siehe Fig. 5B). Bei den IRES-Elementen handelt es sich zum einen um eine der beiden HCV-IRES-Varianten Nukleotide 1-377 oder Nukleotide 1-389, und zum anderen um die IRES des Enzephalomyocarditis Virus, die die Translation der HCV Sequenz strangabwärts von den Genen für NS2 oder NS3 bis zu dem authentischen 3'Ende des 60 Genoms steuert.

Die beiden genannten HCV-IRES-Varianten wurden wie folgt ermittelt:

Auf der Basis von Deletionsanalysen der 3'Grenze der HCV-IRES (Reynolds et al. 1995, EMBO J. 14, 6010) wurden verschiedene Abschnitte der 5'NTR mit dem NPT Gen fusioniert und anhand von Kotransfektionen mit einem das T7 RNA Polymerase Gen enthaltenden Plasmid hinsichtlich der maximalen Anzahl gebildeter Kolonien analysiert. Die besten Ergebnisse wurden mit den HCV Sequenzen von 1-377 und 1-389 erhalten. Da sich das AUG-Startkodon des HCV Polyproteins an Position 342 befindet und somit in der IRES-Sequenz enthalten ist, kommt es zu einer Fusion von 12 bzw. 16 Aminosäuren des HCV-Kapsidproteins ("Core-Proteins") mit der Neomycin Phosphotransferase (siehe Fig. 1A).

Diese modifizierten HCV-Subgenom-Konstrukte erhielten dementsprechend die Bezeichnungen I<sub>377</sub>/NS2-3' (oder



I<sub>377</sub>/NS3-3') und I<sub>389</sub>/NS2-3' (oder I<sub>389</sub>/NS3-3'). Sie sind in Fig. 1A schematisch dargestellt.

Mit in-vitro-Transkripten dieser modifizierten parentalen HCV-Subgenom-Konstrukte I<sub>377</sub>/NS2-3' (oder I<sub>389</sub>/NS3-3) und I<sub>389</sub>/NS2-3' (oder I<sub>389</sub>/NS3-3') wurden verschiedene Zelllinien und Primärzellkulturen von menschlichen Hepatocyten transfiziert.

Als parallele Negativ-Kontrolle zu allen Transfektionsexperimenten wurde zu jedem modifizierten parentalen HCV-Subgenom-Konstrukt ein entsprechend modifiziertes aber defektes Subgenom konstruiert, das sich von dem parentalen dadurch unterscheidet, daß es innerhalb des Leserasters eine Deletion von 10 Aminosäuren aufweist, die das aktive Zentrum der NS5B RNA Polymerase umfaßt (Behrens et al., 1996, EMBO J. 15, 12; und Lohmann et al., 19977 Virol. 71, 8416).

#### (C) Synthese selektierbarer HCV-Genom-Konstrukte

Ein NS2-3'Subgenomkonstrukt, das am 5'Ende mit einem Fragment des Luziferasegens und der vollständigen EMCV-IRES verbunden ist, wurde mit NcoI und SpeI restringiert und mittels präparativer Agarosegelelektrophorese gereinigt. Der so erhaltene Vektor wurde in einer 3-Faktor Ligation mit einem NcoI/NotI-HCV-Fragment, entsprechend den Nukleotidpositionen 342 bis 1968 des HCV-Genoms und mit einem NotI/SpeI-Fragment, entsprechend den Nukleotidpositionen 1968-9605 ligiert. Das entstandene Konstrukt, bei dem das vollständige HCV-Leseraster und die 3'NTR stromabwärts dem Luziferasegenfragment und der EMCV-IRES liegen, wurde danach mit PmeI und SpeI restringiert und mit dem analog restringierten I<sub>389</sub>/NS3-3'/wt-Subgenomkonstrukt-Vektor ligiert. Dieses selektionierbare HCV-Genomkonstrukt ist in Fig. 4 dargestellt.

#### (D) Herstellung von den HCV-RNA-Konstrukten entsprechenden in-vitro-Transkripten

Die vorstehend beschriebenen gereinigten Plasmid DNAs wurden mit ScaI linearisiert und nach Phenol/Chloroform-Extraktion und Isopropanol-Präzipitation in eine In-vitro-Transkriptionsreaktion eingesetzt unter Verwendung der folgenden Komponenten: 80 mM HEPES, pH 7,5, 12,5 mM MgCl<sub>2</sub>, 2 mM Spermidin, 40 mM Dithiothreitol, 2 mM von jedem NTP, 1 Einheit RNasin/µl, 50 µg/ml restringierte DNA und ca. 2 Einheiten/µl T7 RNA Polymerase. Nach 2 Std. bei 37°C wurde die Hälfte der Menge an T7 Polymerase zugegeben und der Reaktionsansatz weitere 2 h inkubiert. Zur Entfernung von DNA wurde die Mischung mit saurem Phenol extrahiert (U. Kedzierski, J. C. Porte, 1991, Bio Techniques 10, 210), mit Isopropanol präzipitiert, das Pellet in Wasser gelöst und mit DNase (2 Einheiten pro µg DNA) für 60 Min. bei 37°C inkubiert. Nach anschließender Extraktion mit saurem Phenol, saurem Phenol/Chloroform und Chloroform und Isopropanol-Präzipitation wurde die gelöste RNA mittel optischer Dichtemessungen quantifiziert und ihre Unversehrtheit mittels Formaldehyd-Agarose-Gelelektrophorese überprüft.

#### Beispiel 2

##### Transfektionsexperimente mit der Hepatomazelllinie Huh-7

Bei sämtlichen Transfektionsexperimenten wurde sorgfältig darauf geachtet, daß jegliche Matrizen-DNA zuvor entfernt worden war, um zu vermeiden, daß solche DNA in transfizierte Zellen integrieren und diesen unabhängig von einer HCV-Replikation eine Neomycin-Resistenz vermitteln konnte. Deshalb wurde im Anschluß an die in-vitro-Transkription (Beispiel 1D) die Reaktionsmischung mit 2 Einheiten DNase pro µg DNA für 60 Min. bei 37°C behandelt und mit saurem Phenol/Chloroform und Chloroform extrahiert. Vor der Verwendung für die Transfektion wurde die präzipitierte RNA mittels Formaldehyd Agarose Gel Elektrophorese analysiert.

Es wurden drei separate Transfektionsexperimente mit der hoch differenzierten humanen Hepatomazelllinie Huh-7 (gemäß Nakabayashi et al. 1982, Cancer Res. 42, 3858) durchgeführt. Dabei wurde jeweils 15 µg RNA in 8 × 10<sup>6</sup> Huh-7 Zellen mit Hilfe der Elektroporation eingebracht und diese Zellen anschließend in Kulturschalen von 10 cm Durchmesser ausgesät. 24 Stunden nach der Aussaat wurde Neomycin (= G418) in einer Endkonzentration von 1 mg/ml zugegeben. Das Kulturmedium wurde zweimal pro Woche gewechselt. Nach 3-5 Wochen waren kleine Kolonien erkennbar, die isoliert und unter den gleichen Kulturbedingungen passagiert wurden.

Die Zellklone, die im Verlauf des ersten Experiments erhalten wurden, wurden isoliert und subpassagiert. Während dieser Prozedur starben die meisten Klone und die Endausbeute betrug nur noch 9 Klone von Zellen, die mit den parentalen HCV-Subgenom-Konstrukten transfiziert worden waren und 1 Klon (Klon 8-1) von Zellen, die mit einem defekten HCV-Genom-Konstrukt, nämlich einer defekten NS2-3'HCV-RNA transfiziert worden waren. Außer einer verkürzten Verdopplungszeit und dem gelegentlichen Auftreten von irregulär geformten Zellen wurden keine beständigen morphologischen Unterschiede zwischen diesen 9 Zellklonen und dem einen Zellklon (Klon 8-1) oder den parentalen Huh-7 Zellen gefunden.

Die Hauptkriterien für funktionierende HCV-Genomkonstrukte sind die Bildung von viraler RNA mit korrekter Größe und die Abwesenheit von (integrierter) Plasmid DNA, die eine G418-Resistenz übertragen bzw. vermitteln könnte.

Um die HCV-RNA in den Huh-7-Zellen zu bestimmen, wurde die Gesamt-RNA isoliert und mittels des gängigen Northern-Blot Verfahrens unter Verwendung einer Plusstrang-spezifischen Ribosonde (= RNA-Sonde) analysiert. Hierfür wurde von den jeweiligen Zellklonen Gesamt-RNA nach der Methode von Chomczynski und Sacchi 1987, Anal. Biochem. 162, 156 isoliert, und 10 µg RNA, was dem Gesamt-RNA-Gehalt von 0,5-1 × 10<sup>6</sup> Zellen entspricht, mittels denaturierender Formaldehyd-Agarose-Gelelektrophorese aufgetrennt (Spuren 3 bis 12 der Fig. 1B). Als Größenmarker mit authentischer Sequenz wurden gleichzeitig 109 in-vitro-Transkripte (ivr.), die zu den I<sub>389</sub>/NS2-3'/wt oder den I<sub>389</sub>/NS3-3'/wt Replikon-RNAs korrespondieren, mit aufgetrennt (Spur 1 bzw. Spur 2). Die aufgetrennte RNA wurde auf Nylon-Membranen transferiert und mit radioaktiv markierter Plusstrang-spezifischer RNA-Sonde, die komplementär zu dem kompletten NPT-Gen und der HCV-IRES von Nukleotid 377 bis Nukleotid 1 war, hybridisiert. Die Positionen der

HCV-spezifischen RNAs (Pfeile) und der 28S rRNA sind rechts von Spur 12 angegeben, die Größen (Anzahlen der Nukleotide) der RNA-Marker sind links von Spur 1 angegeben. Die RNA-Marker-Fragmente enthalten HCV-Sequenzen und hybridisieren deshalb mit der Ribosonde (= RNA-Sonde). Die Ergebnisse dieser Analyse sind in Fig. 1B dargestellt.

Mit Ausnahme des mit dem defekten HCV-Genom-Konstrukt transfizierten Klon 8-1, lieferten alle Zellklone homogene HCV-RNAs korrekter Länge (ca. 8640 Nukleotide im Fall des NS2-3' und ca. 7970 Nukleotide im Fall des NS3-3'Replikons). Dieser Befund ist ein Indiz dafür, daß die funktionalen Replikons bzw. die funktionalen HCV-Genom-Konstrukte die G418 Resistenz übertragen. Um auszuschließen, daß die G418 Resistenz auf eine Plasmid-DNA zurückzuführen ist, die in das Genom der Huh-7 Wirtszelle integriert ist und unter der Kontrolle eines zellulären Promotors transkribiert wird, wurde von jedem Klon die DNA mittels einer NPT-Gen-spezifischen PCR untersucht. Hierbei wurde aus den selektierten Huh-7-Zellklonen die DNA mittels Verdau mit Proteinase K (40 µg/ml, 1 h, 37°C) in 10 mM Tris, pH 7,5, 1 mM EDTA, 0,5% SDS und anschließender Extraktion mit Phenol, Phenol/Chloroform- und Isopropanolpräzipitation isoliert. Das DNA-Präzipitat wurde in 10 mM Tris (pH 7,5) und 1 mM EDTA gelöst und 1 Stunde mit Rnase A inkubiert. Im Anschluß an eine Phenol/Chloroform Extraktion und Ethanol Präzipitation wurde 1 µg DNA, entsprechend 4–8 × 10<sup>4</sup> Zellen, mittels PCR unter Einsatz NPT-Gen-spezifischer Primer (5'-TCAAGACCGACCTG TCCGGTGCCC-3' und 5'-CTTGAGCCTGGCGAACAGTTCGGC-3') analysiert und ein DNA-Fragment bestehend aus 379 Nukleotiden erzeugt. Die Spezifität des PCR-Produkts wurde mittels Southern Blot Verfahren nachgewiesen, wobei ein Digoxigenin-markiertes DNA Fragment eingesetzt wurde, das zu dem NPT-Gen korrespondiert. Als Positiv-Kontrollen (zum Nachweis etwa vorhandener kontaminierender Nukleinsäuren) wurde das PCR-Verfahren mit 10<sup>7</sup> Plasmid Molekülen oder 1 µg DNA aus einer BHK Zelllinie, die stabil mit einem Neomycin-Resistenz-Gen transfiziert war, durchgeführt, und als

Negativ-Kontrolle wurde die PCR mit denselben Reagenzien aber ohne zugesetzte DNA durchgeführt. Die Ergebnisse dieser Untersuchung sind in Fig. 1C dargestellt. Die Spuren 1 und 2 repräsentieren die Positiv-Kontrollen, Spur 13 repräsentiert die Negativ-Kontrolle. Die Zahlenangaben links der Spur 1 bezeichnen die Größe der Nukleotid-Marker-Moleküle. Außer in Klon 7-3 (Fig. 1C, Spur 3), der von Zellen nach Transfektion mit einem NS2-3'-Replikon/NS2-3'HCV-Genom-Konstrukt stammt, und in Klon 8-1 (Fig. 1C, Spur 12), der von Zellen nach Transfektion mit einem defekten HCV-Genom-Konstrukt stammt, war in keinem Zellklon eine NPT-DNA nachweisbar. Dieser Befund ist ein weiteres Indiz dafür, daß die G418 Resistenz der meisten Klone durch die replizierende HCV-RNA vermittelt wurde. Aber auch unabhängig von diesen Ergebnissen ist es unwahrscheinlich, daß HCV-RNAs mit korrekter Größe von integrierter Plasmid DNA erzeugt wird, denn die für die in-vitro-Transkription verwendeten Plasmide enthalten weder einen eukaryontischen Promotor noch ein Polyadenylierungssignal. Im Fall des Klons 7-3 ist die Resistenz deshalb höchst wahrscheinlich sowohl durch das HCV-RNA-Konstrukt bzw. die replizierende HCV-RNA als auch durch eine integrierte NPT-DNA Sequenz vermittelt worden, während die Resistenz der Zellen von Klon 8-1 ausschließlich auf die integrierte Plasmid DNA zurückzuführen ist.

Um zu bestätigen, daß die G418 Resistenz von einer autonom replizierenden HCV-RNA vermittelt ist, wurde der Klon 9-13 (Fig. 1B, Spur 11) weiteren Tests unterworfen. Klon 8-1, der integrierte Kopien des NPT-Gens trägt, wurde überall als Negativkontrolle eingesetzt. Mit dem Ziel, die Anwesenheit von NPT-DNA im Klon 9-13 rigoros auszuschließen, wurde eine PCR durchgeführt, die den Nachweis von < 1000 NPT-Gen-Kopien in 40.000 Zellen erlaubt. Das Ergebnis dieser PCR ist in Fig. 2A dargestellt. Im einzelnen wurde bei dieser PCR wie folgt verfahren:

Es wurden jeweils 10<sup>6</sup>–10<sup>7</sup> Plasmid Moleküle (I<sub>377</sub>/NS3-3'/wt) entweder direkt (Spuren 7–11) oder nach Zugabe von jeweils 1 µg 9-13 DNA (Spuren 2–6) in dem Test eingesetzt. Die Spezifität der amplifizierten DNA Fragment wurde mittels Southern Blot unter Verwendung einer NPT-spezifischen Sonde bestimmt. Eine PCR ohne DNA-Sonde wurde als Negativ-Kontrolle durchgeführt (Spur 12).

Selbst mit dieser sensitiven Methode wurde in einem µg DNA des Zellklons 9-13 keine Plasmid DNA gefunden (Spur 1). Um die Menge an HCV Plus- und Minusstrang RNAs in diesen Zellen abzuschätzen, wurde eine Verdünnungsreihe von Gesamt-RNA mit dem Northern-Blot-Verfahren unter Verwendung einer Plus- oder Minusstrang-spezifischen radioaktiv markierten Ribosonde (= RNA-Sonde) analysiert. Hierfür wurden jeweils 8, 4 oder 2 µg Gesamt-RNA, die aus den Zellklonen 9-13 und 8-1 isoliert worden waren, parallel zu bekannten Mengen analoger in-vitro-Transkripte mit Plus- oder Minusstrang-Polarität (Kontroll-RNAs) im Northern-Blot-Verfahren analysiert und anschließend einer Hybridisierung unterworfen. Die Hybridisierung wurde mit einer Plusstrang-spezifischen Ribosonde, die das komplette NPT-Gen und die HCV-IRES abdeckte ("plusstrang", obere Bildtafel), oder mit einer Minusstrang-spezifischen RNA-Sonde, die zu der NS3-Sequenz komplementär war ("minusstrang", untere Bildtafel) durchgeführt. Die Pfeile markieren die Positionen von Replikon-RNA. Die Ergebnisse dieser Analyse sind in Fig. 2B dargestellt.

Im Fall des Plusstrangs wurden ca. 10<sup>8</sup> Kopien/µg Gesamt-RNA nachgewiesen, was 1000–5000 HCV-RNA-Molekülen pro Zelle entspricht, während die Menge an Minusstrang-RNA 5- bis 10-fach niedriger war. Dieses Ergebnis stimmt mit der Annahme überein, daß die Minusstrang RNA die replikative Zwischenform bzw. Zwischenkopie ist, die als Vorlage für die Synthese der Plusstrang Moleküle dient.

Da die Reaktion im wesentlichen von der viralen RNA-abhängigen RNA Polymerase katalysiert wird, sollte die Synthese der HCV-RNAs resistent gegen Dactinomycin sein, einem Antibiotikum, das selektiv die RNA-Synthese von DNA-Matrizen inhibiert, nicht jedoch die RNA-Synthese von RNA-Matrizen. Um diese Vermutung zu bestätigen, wurden Zellen mit [<sup>3</sup>H] Uridin in Anwesenheit von Dactinomycin inkubiert, die radioaktiv markierten RNAs extrahiert, mittels denaturierender Agarose-Gel-Elektrophorese aufgetrennt und mit Hilfe eines handelsüblichen Bio-Imagers unter Verwendung einer [<sup>3</sup>H]-sensitiven Bildplatte analysiert. Hierfür wurden jeweils ca. 5 × 10<sup>5</sup> Zellen der Klone 9-13 und 8-1 mit 100 µCi [<sup>3</sup>H]Uridin für 16 Std. in Abwesenheit (–) oder Gegenwart (+) von 4 µg/ml Dactinomycin (Dact) inkubiert. Im Anschluß an diese Markierungsreaktion wurde die Gesamt-RNA präpariert und mittels Formaldehyd-Agarose-Gel-Elektrophorese analysiert. In den beiden ersten Spuren ist nur 1/10 der Gesamt-RNA dargestellt. Die radioaktiv markierte RNA wurden mit einem BAS-2500 Bio-Imager (Firma Fuji) sichtbar gemacht.

Die Ergebnisse dieser Analyse sind in Fig. 2C dargestellt. In Übereinstimmung mit dem Inhibitor-Profil der NS5B Polymerase (Behrens et al., 1996, EMBOJ. 15, 12 und Lohmann et al., 1997, J. Virol. 71, 8416) war die Replikation der HCV-RNA nicht durch Dactinomycin beeinflusst worden, während die Synthese von zellulärer RNA gehemmt worden

war. Um die Identität der viralen RNA zu bestätigen, wurde eine RT-PCR zur Reklonierung der replizierten Sequenzen durchgeführt. Die Sequenzanalyse der reklonierten RNA zeigte, daß die RNA in dem Klon 9-13 HCV-spezifisch ist und mit dem transfizierten Transkript des HCV-Konstrukts I<sub>377</sub>/NS3-3'/wt übereinstimmt.

Zur Analyse der viralen Proteine wurden die betreffenden Zellen zunächst metabolisch mit [<sup>35</sup>S] Methionin/Cystein radioaktiv markiert, anschließend lysiert und danach die HCV-spezifischen Proteine mittels Immunopräzipitation aus den Zell-Lysaten isoliert. Die Ergebnisse dieser Analysen sind in Fig. 3 A dargestellt. Im einzelnen wurde dabei wie folgt verfahren: Zellen der Zellklone 9-13 (wt) und 8-1 (Δ) waren durch Behandlung für 16 Stunden mit einer dem Fachmann geläufigen und im Handel erhältlichen Protein-Markierungs-Mischung (z. B. NEN Life Science) metabolisch radioaktiv markiert worden. Mittels Immunopräzipitation (IP) unter nichtdenaturierenden Bedingungen (z. B. nach Bartenschlager et al., 1995, J. Virol. 69, 7519) und unter Verwendung von drei verschiedenen Antisera (3/4, 5A, 5B, gemäß Markierung am oberen Ende der Spuren 1 bis 12) waren die HCV-spezifischen Proteine vom Zell-Lysat abgetrennt worden. Die Immunkomplexe wurden mittels Tricine SDS-PAGE analysiert und mittels Autoradiographie sichtbar gemacht. Um authentische Größenmarker zu erhalten, wurde das homologe Replikonstrukt I<sub>377</sub>/NS3-3'/wt einer transienten Expression mit dem Vaccinia Virus T7-Hybrid System in Huh-7 Zellen unterworfen. Die dabei erhaltenen Produkte waren als Größenmarker (Spuren 7-9) parallel zu den Zellen der Klone 9-13 und 8-1 behandelt worden. Identifizierte HCV-Proteine sind am linken Rand von Spur 1 markiert, die Molekulargewichte (in Kilodalton) sind am rechten Rand von Spur 9 angegeben. Es ist anzumerken, daß das verwendete NS3/4-spezifische Antiserum ("3/4") bevorzugt mit NS4A und NS4B reagiert, was zu einer Unterrepräsentation von NS3 führt.

Alle viralen Antigene waren eindeutig nachweisbar und ihre apparenten Molekulargewichte zeigten keine Abweichungen gegenüber denjenigen, die nach transients Expression desselben bicistronischen HCV-RNA-Konstrukts in den ursprünglichen Huh-7 Zellen ermittelt wurden. Um die subzelluläre Verteilung der viralen Antigene zu bestimmen, wurde eine Immunofluoreszenz-Nachweisreaktion unter Einsatz von NS3- und NS5A-spezifischen Antisera durchgeführt (z. B. nach Bartenschlager et al., 1995, J. Virol. 69, 7519). Hierfür wurden Zellen der Klone 9-13 (wt) und 8-1 (Δ) 24 Std. nach dem Aussäen auf Deckgläsern mit Methanol/Azeton fixiert und mit polyklonalen NS3- oder NS5A-spezifischen Antisera inkubiert. Die gebundenen Antikörper wurden mit einem kommerziell erhältlichen FITC-konjugierten Anti-Kaninchen-Antiserum sichtbar gemacht. Zur Unterdrückung unspezifischer Fluoreszenzsignale wurden die Zellen mit dem Farbstoff "Evans Blue" gegengefärbt.

Die Ergebnisse dieses Nachweistests sind in Fig. 3 B dargestellt. Mit beiden Antisera war eine starke Fluoreszenz im Zytoplasma nachweisbar. Die NS5A-spezifischen Antisera führten außerdem zu einer schwachen Zellkern-Fluoreszenz, was darauf hindeutet, daß zumindest kleine Mengen dieses Antigens auch zum Zellkern gelangen. Die generell dominierende Präsenz der viralen Antigene im Zytoplasma ist jedoch ein starkes Indiz dafür, daß die HCV-RNA Replikation im Zytoplasma stattfindet – so wie das bei den meisten RNA-Viren der Fall ist.

Diese Ergebnisse belegen klar, daß mit dem hier beschriebenen Versuchsansatz der Aufbau eines Zellkultursystems für das HCV gelungen ist, dessen Effizienz alles bisher bekannte um Größenordnungen übersteigt und erstmalig den Nachweis viraler Nukleinsäuren und Proteine mit konventionellen und bewährten biochemischen Methoden erlaubt. Erst diese Effizienz erlaubt überhaupt detaillierte Untersuchungen der HCV-Pathogenese, genetische Analysen verschiedener HCV-Funktionen und ein genaues Studium der Virus-/Wirtszellwechselwirkungen, wodurch sich neue Ansatzpunkte für die Entwicklung einer antiviralen Therapie definieren lassen.

#### Beispiel 3

##### Transfektion von Huh-7 Zellen mit HCV-Genomkonstrukten

Huh-7 Zellen werden wie in Beispiel 2 beschrieben transfiziert und selektioniert, wobei hier jedoch selektionierbare Konstrukte verwendet werden, die das vollständige Virusgenom enthalten. Die erhaltenen Zellklone werden analog dem Beispiel 2 mittels PCR auf Abwesenheit von HCV-DNA untersucht und die produktive Replikation der HCV-RNA wird danach mittels Northern Blot, [<sup>3</sup>H]Uridinmarkierung in Anwesenheit von Dactinomycin, Nachweis der viralen Proteine bzw. Antigene vorzugsweise mit Hilfe des Western Blots, der Immunopräzipitation oder der Immunfluoreszenz nachgewiesen. Im Gegensatz zu den im Beispiel 2 beschriebenen Ansätzen lassen sich mit dem hier beschriebenen Konstrukt außerdem vollständige und sehr wahrscheinlich infektiöse Viren erhalten, was bei den dort (in Beispiel 2) beschriebenen Subgenomkonstrukten nicht der Fall ist. Diese Viren, die in der Zelle und dem Zellkulturüberstand vorhanden sind, werden beispielsweise mittels Ultrazentrifugation, Immunopräzipitation oder Fällung mit Polyethylenglykol konzentriert und alle exogenen, d. h. nicht im Viruspartikel eingebauten Nukleinsäuren werden mittels Inkubation mit Nukleasen (RNase, DNase, Mikrococcusnuklease) verdaut. Auf diese Weise lassen sich alle kontaminierenden Nukleinsäuren, die nicht im schützenden Viruspartikel enthalten sind, entfernen. Die geschützte virale RNA wird nach Inaktivierung der Nukleasen, beispielsweise mittels Inkubation mit Proteinase K in einem SDS-haltigen Puffer durch Extraktion mit Phenol und Phenol/Chloroform isoliert und mittels Northern Blot oder RT-PCR unter Verwendung HCV-spezifischer Primer nachgewiesen. Auch in diesem Versuchsansatz ist die Kombination des beschriebenen HCV-Konsensusgenoms mit einem Selektionsmarker entscheidend für die effiziente Produktion von viraler RNA, viralem Protein und damit von HCV-Partikeln.

#### Beispiel 4

Herstellung und Anwendung eines HCV-RNA Konstrukts, bei dem das Resistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-Subgenom-Sequenz verbunden ist

Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1 oder Beispiel 3 hergestellt, bei dem ein Antibiotikumresistenzgen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA-Sequenz verbunden ist. Solche Konstrukte sind in Fig. 7 schematisch dargestellt. Huh-7 Zellen werden wie in Beispiel 2 beschrieben mit diesem HCV-RNA-

Konstrukt transfiziert. Nach der Transfektion in die Zellen erfolgt zunächst die Selektion mit dem entsprechenden Antibiotikum. In den dabei erhaltenen Zellklonen wird das einklonierte Ribozym aktiviert oder, im Fall eines Konstrukts, das eine Erkennungsstelle für ein Ribozym trägt, wird das Ribozym in die Zelle eingeschleust (z. B. mittels Transfektion eines Ribozymkonstrukts oder Infektion mit einem viralen Expressionsvektor, in den das entsprechende Ribozym eingesetzt wurde). In beiden Fällen wird durch die ribozymvermittelte Spaltung das Resistenzgen von der HCV-RNA-Sequenz abgetrennt. Das Ergebnis ist im Fall des HCV-Genom-Konstrukts ein authentisches HCV-Genom ohne Resistenzgen, das zur Bildung authentischer infektiöser Viruspartikel befähigt ist. Im Fall des HCV-Subgenom-Konstrukts entsteht ein HCV-Replikon ohne Resistenzgen.

#### Beispiel 5

##### Kotransfektion eines HCV-RNA-Konstrukts mit einem separaten Luziferase-Transfektionskonstrukt

Es wird ein HCV-RNA-Konstrukt gemäß Beispiel 1(A) oder Beispiel 3 oder Beispiel 4 hergestellt. Parallel dazu wird ein Transfektionskonstrukt hergestellt, welches das Luziferasegen umfaßt, wobei dieses Luziferasegen mittels einer ersten Nukleotidsequenz, die für eine HCV-Protease- (z. B. NS3-Protease-)Spaltungsstelle kodiert, mit einer zweiten Nukleotidsequenz, die für ein anderes Protein oder einen Teil eines anderen Proteins kodiert, verbunden ist. HCV-RNA-Konstrukt und Transfektionskonstrukt werden in beliebige Wirtszellen, vorzugsweise Hepatomazellen, insbesondere Huh-7-Zellen, eingeschleust. Dies kann auf die in Beispiel 2 beschriebene Art und Weise geschehen. Das Produkt des modifizierten Luziferasegens ist ein Luziferase-Fusionsprotein, in dem die Luziferase auf Grund der Fusion mit dem Fremdanteil inaktiv ist. In transfizierten Zellen mit hoher HCV-Replikation wird das Fusionsprotein, das ja eine Schnittstelle für eine HCV-Protease enthält, gespalten und damit die aktive Form der Luziferase freigesetzt, die sich durch luminometrische Messung bestimmen läßt. Wird die Replikation des HCV-RNA-Konstrukts gehemmt, wird das Fusionsprotein nicht gespalten und keine aktive Luziferase freigesetzt. Infolgedessen ist die quantitative Bestimmung der Luziferase ein Maß für die Replikation des HCV-Subgenom-Konstrukts. Anstelle des Luziferasegens kann ebenso ein anderes Reportergen verwendet werden, das in analoger Weise modifiziert ist, so daß seine Expression von der Virusreplikation abhängt, obwohl dieses Reportergen nicht Bestandteil des HCV-Subgenom-Konstrukts ist. Es kann auch ein zelluläres Protein, welches durch die HCV-Proteine oder Nukleinsäure inaktiviert oder aktiviert wird, als sogenannter Surrogatmarker verwendet werden. In diesem Fall ist die Expression bzw. Aktivität dieses Surrogatmarkers ein Maß für die Replikation der viralen DNA.

#### Beispiel 6

##### Herstellung von HCV-Subgenom-Konstrukten mit integrierten Fremdgenen zur Verwendung als leberzellspezifische Genfähren für die Gentherapie

Diese rekombinanten und selektionierbaren HCV-Subgenom-Konstrukte werden in trans-komplementierende Helferzelllinien transfiziert, d. h. in Zelllinien, die induzierbar oder konstitutiv die fehlenden Funktionen (beispielsweise die Strukturproteine) exprimieren. Zellklone, die ein funktionelles HCV-Subgenom-Konstrukt enthalten, lassen sich durch entsprechende Selektion etablieren. Die von der Wirtszelle exprimierten Virus-Strukturproteine erlauben die Bildung von Viruspartikeln, in die die RNA der HCV-Subgenom-Konstrukte eingeschleust wird. Das Ergebnis sind also virusähnliche Partikel, die ein erfindungsgemäßes HCV-Subgenom-Konstrukt einschließlich des einklonierten Fremdgens enthalten und die dieses mittels Infektion auf andere Zellen übertragen können. Ein Beispiel für ein solches Konstrukt ist in Fig. 8 dargestellt. Es besteht auch die Möglichkeit, das hier beschriebene erfindungsgemäße HCV-Subgenom-Konstrukt mit integriertem Fremdgen direkt als Expressionsvektor einzusetzen. Dabei wird analog dem vorgenannten Verfahren vorgegangen, allerdings mit dem Unterschied, daß Zelllinien transfiziert werden, die keine transkomplementierenden Faktoren exprimieren. In diesem Fall dient das HCV-Konstrukt also lediglich als Expressionsvektor.

# DE 199 15 178 A 1

## SEQUENZPROTOKOLL

### ALLGEMEINE ANGABEN:

#### ANMELDER:

NAME: Universität Mainz  
STRASSE: Obere Zahlbacher Str. 67...  
ORT: Mainz  
BUNDESLAND: Rheinland-Pfalz  
LAND: Deutschland  
POSTLEITZAHL: 55101

#### VERTRETER:

NAME: Dr. Ulrike Rudolph  
STRASSE: In der Schanz 10  
ORT: Schriesheim  
BUNDESLAND: Baden-Württemberg  
LAND: Deutschland  
POSTLEITZAHL: 69198  
VERTRETERNUMMER: 246 263  
AKTENZEICHEN: .....

#### TELEKOMMUNIKATION:

TELEFON: 06203-61348  
TELEFAX: 06203-64196

### BEZEICHNUNG DER ERFINDUNG:

Hepatitis C Virus Zellkultursystem

### ANZAHL DER SEQUENZEN: 5

### COMPUTERLESBARE FASSUNG:

DATENTRÄGER: Diskette  
COMPUTER: PC  
BETRIEBSSYSTEM: MS DOS  
SOFTWARE: Word 6.0 für windows 95

# DE 199 15 178 A 1

## ANGABEN ZU SEQ ID-NO:1:

5 SEQUENZCHARAKTERISTIKA:  
LÄNGE: 11076 Basenpaare  
ART: Nukleinsäure  
STRANGFORM: Einzelstrang  
10 TOPOLOGIE: linear

ART DES MOLEKÜLS: Genom-RNA

15 HYPOTHETISCH: Nein

ANTI-SENSE: nein

20 URSPRÜNGLICHE HERKUNFT:  
ORGANISMUS: Hepatitis C Virus  
STAMM: Genotyp 1b  
25 ENTWICKLUNGSSTADIUM: —  
ZELLTYP: —

UNMITTELBARE HERKUNFT:  
30 Isolat aus humanen Leberzellen

MERKMAL:  
NAME/SCHLÜSSEL: kodierende Sequenz für Viruspartikel  
35 LAGE: —  
ERMITTLUNGSMETHODE: . experimentell

## SEQUENZBESCHREIBUNG: SEQ ID-NO: 1

45 I<sub>389</sub>/Core-3'/wt: 11076 bp

50	GCCAGCCCCC GATTGGGGGC GAACTCCAC CATAGATCAC TCCCCTGTGA	50
	GGAAGTACTG TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG	100
	TGTCGTGCAG CCTCCAGGAC CCCCCCTCCC GGGAGAGCCA TAGTGGTCTG	150
55	CGGAACCGGT GAGTACACCG GAATTGCCAG GACGACCGGG TCCTTTCTTG	200
	GATCAACCCG CTCAATGCCT GGAGATTGCG GCGTGCCCCC GCGAGACTGC	250
60	TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG	300
	GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCACG	350
65	AATCCTAAAC CTCAAAGAAA AACCACACGT AACACCAACG GGCGCGCCAT	400
	GATTGAACAA GATGGATTGC ACGCAGGTTT TCCGGCCGCT TGGGTGGAGA	450

# DE 199 15 178 A 1

GGCTATTTCGG CTATGACTGG GCACAACAGA CAATCGGCTG CTCTGATGCC	500	
GCCGTGTTCC GGCTGTCAGC GCAGGGGGCG CCGGTTCTTT TTGTCAAGAC	550	
CGACCTGTCC GGTGCCCTGA ATGAACTGCA GGACGAGGCA GCGCGGCTAT	600	5
CGTGGCTGGC CACGACGGGC GTTCCTTGCG CAGCTGTGCT CGACGTTGTC	650	
ACTGAAGCGG GAAGGGACTG GCTGCTATTG GGCGAAGTGC CGGGGCAGGA	700	10
TCTCCTGTCA TCTCACCTTG CTCCTGCCGA GAAAGTATCC ATCATGGCTG	750	
ATGCAATGCG GCGGCTGCAT ACGCTTGATC CGGCTACCTG CCCATTCGAC	800	15
CACCAAGCGA AACATCGCAT CGAGCGAGCA CGTACTCGGA TGGAAAGCCGG	850	
TCTTGTCGAT CAGGATGATC TGGACGAAGA GCATCAGGGG CTCGCGCCAG	900	20
CCGAACTGTT CGCCAGGCTC AAGGCGCGCA TGCCCGACGG CGAGGATCTC	950	
GTCGTGACCC ATGGCGATGC CTGCTTGCCG AATATCATGG TGGAAAATGG	1000	25
CCGCTTTTCT GGATTCATCG ACTGTGGCCG GCTGGGTGTG GCGGACCGCT	1050	
ATCAGGACAT AGCGTTGGCT ACCCGTGATA TTGCTGAAGA GCTTGGCGGC	1100	
GAATGGGCTG ACCGCTTCCT CGTGCTTTAC GGTATCGCCG CTCCCGATTC	1150	30
GCAGCGCATC GCCTTCTATC GCCTTCTTGA CGAGTTCTTC TGAGTTTAAA	1200	
CAGACCACAA CGGTTTCCCT CTAGCGGGAT CAATTCCGCC CCTCTCCCTC	1250	35
CCCCCCCCCT AACGTTACTG GCCGAAGCCG CTTGGAATAA GGCCGGTGTG	1300	
CGTTTGTCTA TATGTTATTT TCCACCATAT TGCCGTCTTT TGGCAATGTG	1350	40
AGGGCCCGGA AACCTGGCCC TGTCTTCTTG ACGAGCATTC CTAGGGGTCT	1400	
TTCCCCTCTC GCCAAAGGAA TGCAAGGTCT GTTGAATGTC GTGAAGGAAG	1450	45
CAGTTCCTCT GGAAGCTTCT TGAAGACAAA CAACGTCTGT AGCGACCCTT	1500	
TGCAGGCAGC GGAACCCCCC ACCTGGCGAC AGGTGCCTCT GCGGCCAAAA	1550	
GCCACGTGTA TAAGATACAC CTGCAAAGGC GGCACAACCC CAGTGCCACG	1600	50
TTGTGAGTTG GATAGTTGTG GAAAGAGTCA AATGGCTCTC CTCAAGCGTA	1650	
TTCAACAAGG GGCTGAAGGA TGCCCAGAAG GTACCCCATC GTATGGGATC	1700	55
TGATCTGGGG CCTCGGTGCA CATGCTTTAC ATGTGTTTAG TCGAGGTAA	1750	
AAAACGTCTA GGCCCCCGA ACCACGGGGA CGTGGTTTTT CTTTGAAAAA	1800	60
CACGATAATA CCATGGGCAC GAATCCTAAA CCTCAAAGAA AAACCAAACG	1850	
TAACACCAAC CGCCGCCAC AGGACGTCAA GTTCCCGGGC GGTGGTCAGA	1900	
TCGTCGGTGG AGTTTACCTG TTGCCGCGCA GGGGCCCCAG GTTGGGTGTG	1950	65

# DE 199 15 178 A 1

	CGCGCGACTA	GGAAGACTTC	CGAGCGGTCG	CAACCTCGTG	GAAGGCGACA	2000
	ACCTATCCCC	AAGGCTCGCC	AGCCCGAGGG	TAGGGCCTGG	GCTCAGCCCG	2050
5	GGTACCCCTG	CCCCCTCTAT	GGCAATGAGG	GCTTGGGGTG	GGCAGGATGG	2100
	CTCCTGTCAC	CCCGTGGCTC	TCGGCCTAGT	TGGGGCCCCA	CGGACCCCG	2150
10	GCGTAGGTCG	CGCAATTTGG	GTAAGGTCAT	CGATACCCTC	ACGTGCGGCT	2200
	TCGCCGATCT	CATGGGGTAC	ATTCCGCTCG	TCGGCGCCCC	CCTAGGGGGC	2250
15	GCTGCCAGGG	CCCTGGCGCA	TGGCGTCCGG	GTTCTGGAGG	ACGGCGTGAA	2300
	CTATGCAACA	GGGAATCTGC	CCGGTTGCTC	CTTTTCTATC	TTCCTTTTGG	2350
20	CTTTGCTGTC	CTGTTTGACC	ATCCCAGCTT	CCGCTTATGA	AGTGCGCAAC	2400
	GTATCCGGAG	TGTACCATGT	CACGAACGAC	TGCTCCAACG	CAAGCATTTG	2450
	GTATGAGGCA	GCGGACATGA	TCATGCATAC	CCCCGGGTGC	GTGCCCTGCG	2500
25	TTCGGGAGAA	CAACTCCTCC	CGCTGCTGGG	TAGCGCTCAC	TCCCACGCTC	2550
	GCGGCCAGGA	ACGCTAGCGT	CCCCACTACG	ACGATACGAC	GCCATGTCGA	2600
30	TTTGCTCGTT	GGGGCGGCTG	CTCTCTGCTC	CGCTATGTAC	GTGGGAGATC	2650
	TCTGCGGATC	TGTTTTCTC	GTCGCCCAGC	TGTTACCTT	CTCGCCTCGC	2700
35	CGGCACGAGA	CAGTACAGGA	CTGCAATTGC	TCAATATATC	CCGGCCACGT	2750
	GACAGGTCAC	CGTATGGCTT	GGGATATGAT	GATGAACTGG	TCACCTACAG	2800
40	CAGCCCTAGT	GGTATCGCAG	TTACTCCGGA	TCCCACAAGC	TGTCGTGGAT	2850
	ATGGTGGCGG	GGGCCCATTG	GGGAGTCCTA	GCGGGCCTTG	CCTACTATTC	2900
	CATGGTGGGG	AACTGGGCTA	AGGTTCTGAT	TGTGATGCTA	CTCTTTGCCG	2950
45	GCGTTGACGG	GGGAACCTAT	GTGACAGGGG	GGACGATGGC	CAAAAACACC	3000
	CTCGGGATTA	CGTCCCTCTT	TTCACCCGGG	TCATCCCAGA	AAATCCAGCT	3050
50	TGTAAACACC	AACGGCAGCT	GGCACATCAA	CAGGACTGCC	CTGAACTGCA	3100
	ATGACTCCCT	CAACACTGGG	TTCTTGCTG	CGCTGTTCTA	CGTGCACAAG	3150
55	TTCAACTCAT	CTGGATGCCC	AGAGCGCATG	GCCAGCTGCA	GCCCCATCGA	3200
	CGCGTTCGCT	CAGGGGTGGG	GGCCCATCAC	TTACAATGAG	TCACACAGCT	3250
	CGGACCAGAG	GCCTTATTGT	TGGCACTACG	CACCCCGGCC	GTGCGGTATC	3300
60	GTACCCGCGG	CGCAGGTGTG	TGGTCCAGTG	TACTGCTTCA	CCCCAAGCCC	3350
	TGTCGTGGTG	GGGACGACCG	ACCGGTTTCG	CGTCCCTACG	TACAGTTGGG	3400
65	GGGAGAATGA	GACGGACGTG	CTGCTTCTTA	ACAACACGCG	GCCGCCGCAA	3450



# DE 199 15 178 A 1

GGCAACTGGT TTGGCTGTAC ATGGATGAAT AGCACTGGGT TCACCAAGAC	3500	
GTGCGGGGGC CCCCCGTGTA ACATCGGGGG GATCGGCAAT AAAACCTTGA	3550	
CCTGCCCCAC GGA CTGCTTC CGGAAGCACC CCGAGGCCAC TTACACCAAG	3600	5
TGTGGTTCGG GGCCTTGGTT GACACCCAGA TGCTTGGTCC ACTACCCATA	3650	
CAGGCTTTGG CACTACCCCT GCACTGTCAA CTTTACCATC TTCAAGGTTA	3700	10
GGATGTACGT GGGGGGAGTG GAGCACAGGC TCGAAGCCGC ATGCAATTGG	3750	
ACTCGAGGAG AGCGTTGTAA CCTGGAGGAC AGGGACAGAT CAGAGCTTAG	3800	15
CCCGCTGCTG CTGTCTACAA CGGAGTGGCA GGTATTGCCC TGTTCCTTCA	3850	
CCACCCTACC GGCTCTGTCC ACTGGTTTGA TCCATCTCCA TCAGAACGTC	3900	20
GTGGACGTAC AATACCTGTA CGGTATAGGG TCGGCGGTTG TCTCCTTTGC	3950	
AATCAAATGG GAGTATGTCC TGTGCTCTT CCTTCTTCTG GCGGACGCGC	4000	25
GCGTCTGTGC CTGCTTGTGG ATGATGCTGC TGATAGCTCA AGCTGAGGCC	4050	
GCCCTAGAGA ACCTGGTGGT CCTCAACGCG GCATCCGTGG CCGGGGCGCA	4100	
TGGCATTCTC TCCTTCCTCG TGTTCCTCTG TGCTGCCTGG TACATCAAGG	4150	30
GCAGGCTGGT CCCTGGGGCG GCATATGCCC TCTACGGCGT ATGGCCGCTA	4200	
CTCCTGCTCC TGCTGGCGTT ACCACCACGA GCATACGCCA TGGACCGGGA	4250	35
GATGGCAGCA TCGTGCGGAG GCGCGGTTT CGTAGGTCTG ATACTCTTGA	4300	
CCTTGTCACC GCACTATAAG CTGTTCCCTCG CTAGGCTCAT ATGGTGGTTA	4350	40
CAATATTTTA TCACCAGGGC CGAGGCACAC TTGCAAGTGT GGATCCCCC	4400	
CCTCAACGTT CGGGGGGGCC GCGATGCCGT CATCCTCCTC ACGTGCGCGA	4450	
TCCACCCAGA GCTAATCTTT ACCATCACCA AAATCTTGCT CGCCATACTC	4500	45
GGTCCACTCA TGGTGCTCCA GGCTGGTATA ACCAAAGTGC CGTACTTCGT	4550	
GCGCGCACAC GGGCTCATTC GTGCATGCAT GCTGGTGCGG AAGGTTGCTG	4600	50
GGGGTCATTA TGTCCAAATG GCTCTCATGA AGTTGGCCGC ACTGACAGGT	4650	
ACGTACGTTT ATGACCATCT CACCCCACTG CGGGACTGGG CCCACGCGGG	4700	55
CCTACGAGAC CTTGCGGTGG CAGTTGAGCC CGTCGTCTTC TCTGATATGG	4750	
AGACCAAGGT TATCACCTGG GGGGCAGACA CCGCGGCGTG TGGGGACATC	4800	60
ATCTTGGGCC TGCCCGTCTC CGCCCGCAGG GGGAGGGAGA TACATCTGGG	4850	
ACCGGCAGAC AGCCTTGAAG GGCAGGGGTG GCGACTCCTC GCGCCTATTA	4900	
CGGCCTACTC CCAACAGACG CGAGGCCTAC TTGGCTGCAT CATCACTAGC	4950	65

# DE 199 15 178 A 1

CTCACAGGCC GGGACAGGAA CCAGGTCGAG GGGGAGGTCC AAGTGGTCTC 5000  
 CACCGCAACA CAATCTTTCC TGGCGACCTG CGTCAATGGC GTGTGTTGGA 5050  
 5 CTGTCTATCA TGGTGCCGGC TCAAAGACCC TTGCCGGCCC AAAGGGCCCA 5100  
 ATCACCCAAA TGTACACCAA TGTGGACCAG GACCTCGTCG GCTGGCAAGC 5150  
 10 GCCCCCGGG GCGCGTTCCT TGACACCATG CACCTGCGGC AGCTCGGACC 5200  
 TTTACTTGGT CACGAGGCAT GCCGATGTCA TTCCGGTGCG CCGGCGGGGC 5250  
 15 GACAGCAGGG GGAGCCTACT CTCCCCCAGG CCCGTCTCCT ACTTGAAGGG 5300  
 CTCTTCGGGC GGTCCACTGC TCTGCCCCC GGGGCACGCT GTGGGCATCT 5350  
 20 TTCGGGCTGC CGTGTGCACC CGAGGGGTTG CGAAGGCGGT GGACTTTGTA 5400  
 CCCGTCGAGT CTATGGAAAC CACTATGCGG TCCCCGGTCT TCACGGACAA 5450  
 CTCGTCCCCT CCGGCCGTAC CGCAGACATT CCAGGTGGCC CATCTACACG 5500  
 25 CCCCTACTGG TAGCGGCAAG AGCACTAAGG TGCCGGCTGC GTATGCAGCC 5550  
 CAAGGGTATA AGGTGCTTGT CCTGAACCCG TCCGTCGCCG CCACCCTAGG 5600  
 30 TTTCCGGGCG TATATGTCTA AGGCACATGG TATCGACCCT AACATCAGAA 5650  
 CCGGGGTAAG GACCATCACC ACGGGTGCCC CCATCACGTA CTCCACCTAT 5700  
 35 GGCAAGTTTC TTGCCGACGG TGTTTGCTCT GGGGGCGCCT ATGACATCAT 5750  
 AATATGTGAT GAGTGCCACT CAACTGACTC GACCACTATC CTGGGCATCG 5800  
 GCACAGTCCT GGACCAAGCG GAGACGGCTG GAGCGCGACT CGTCGTGCTC 5850  
 40 GCCACCGCTA CGCCTCCGGG ATCGGTCACC GTGCCACATC CAAACATCGA 5900  
 GGAGGTGGCT CTGTCCAGCA CTGGAGAAAT CCCCTTTTAT GGCAAAGCCA 5950  
 45 TCCCCATCGA GACCATCAAG GGGGGGAGGC ACCTCATTTT CTGCCATTCC 6000  
 AAGAAGAAAT GTGATGAGCT CGCCGCGAAG CTGTCCGGCC TCGGACTCAA 6050  
 50 TGCTGTAGCA TATTACCGGG GCCTTGATGT ATCCGTCATA CCAACTAGCG 6100  
 GAGACGTCAT TGTCGTAGCA ACGGACGCTC TAATGACGGG CTTTACCGGC 6150  
 GATTTCGACT CAGTGATCGA CTGCAATACA TGTGTCACCC AGACAGTCGA 6200  
 55 CTTACGCTG GACCCGACCT TCACCATTGA GACGACGACC GTGCCACAAG 6250  
 ACGCGGTGTC ACGCTCGCAG CGGCGAGGCA GGA CTGGTAG GGGCAGGATG 6300  
 60 GGCATTTACA GGT TTGTGAC TCCAGGAGAA CGGCCCTCGG GCATGTTCTGA 6350  
 TTCCTCGGTT CTGTGCGAGT GCTATGACGC GGGCTGTGCT TGGTACGAGC 6400  
 65 TCACGCCCGC CGAGACCTCA GTTAGGTTGC GGGCTTACCT AAACACACCA 6450

# DE 199 15 178 A 1

GGGTTGCCCCG TCTGCCAGGA CCATCTGGAG TTCTGGGAGA GCGTCTTTAC	6500	
AGGCCTCACC CACATAGACG CCCATTTCTT GTCCCAGACT AAGCAGGCAG	6550	
GAGACAACTT CCCCTACCTG GTAGCATACC AGGCTACGGT GTGCGCCAGG	6600	5
GCTCAGGCTC CACCTCCATC GTGGGACCAA ATGTGGAAGT GTCTCATACTG	6650	
GCTAAAGCCT ACGCTGCACG GGCCAACGCC CCTGCTGTAT AGGCTGGGAG	6700	10
CCGTTCAAAA CGAGGTTACT ACCACACACC CCATAACCAA ATACATCATG	6750	
GCATGCATGT CGGCTGACCT GGAGGTCGTC ACGAGCACCT GGGTGCTGGT	6800	15
AGGCGGAGTC CTAGCAGCTC TGGCCGCGTA TTGCCTGACA ACAGGCAGCG	6850	
TGGTCATTGT GGGCAGGATC ATCTTGTCGG GAAAGCCGGC CATCATTCCC	6900	20
GACAGGGAAG TCCTTTACCG GGAGTTCGAT GAGATGGAAG AGTGCGCCTC	6950	
ACACCTCCCT TACATCGAAC AGGGAATGCA GCTCGCCGAA CAATTCAAAC	7000	25
AGAAGGCAAT CGGGTTGCTG CAAACAGCCA CCAAGCAAGC GGAGGCTGCT	7050	
GCTCCCGTGG TGAATCCAA GTGGCGGACC CTCGAAGCCT TCTGGGCGAA	7100	
GCATATGTGG AATTTTCATCA GCGGGATACA ATATTTAGCA GGCTTGTCCTA	7150	30
CTCTGCCTGG CAACCCCGCG ATAGCATCAC TGATGGCATT CACAGCCTCT	7200	
ATCACCAGCC CGCTCACCAC CCAACATACC CTCCTGTTTA ACATCCTGGG	7250	35
GGGATGGGTG GCCGCCCAAC TTGCTCCTCC CAGCGCTGCT TCTGCTTTTCG	7300	
TAGGCGCCGG CATCGCTGGA GCGGCTGTTG GCAGCATAGG CCTTGGAAG	7350	40
GTGCTTGTTG ATATTTTGGC AGGTTATGGA GCAGGGGTGG CAGGCGCGCT	7400	
CGTGGCCTTT AAGGTCATGA GCGGCGAGAT GCCCTCCACC GAGGACCTGG	7450	
TTAACCTACT CCCTGCTATC CTCTCCCTG GCGCCCTAGT CGTCGGGGTC	7500	45
GTGTGCGCAG CGATACTGCG TCGGCACGTG GGCCAGGGG AGGGGGCTGT	7550	
GCAGTGGATG AACCGGCTGA TAGCGTTCGC TTCGCGGGGT AACCACGTCT	7600	50
CCCCACGCA CTATGTGCCT GAGAGCGACG CTGCAGCACG TGTCACCTCAG	7650	
ATCCTCTCTA GTCTTACCAT CACTCAGCTG CTGAAGAGGC TTCACCACTG	7700	55
GATCAACGAG GACTGCTCCA CGCCATGCTC CGGCTCGTGG CTAAGAGATG	7750	
TTTGGGATTG GATATGCACG GTGTTGACTG ATTTCAAGAC CTGGCTCCAG	7800	60
TCCAAGCTCC TGCCGCGATT GCCGGGAGTC CCCTTCTTCT CATGTCAACG	7850	
TGGGTACAAG GGAGTCTGGC GGGGCGACGG CATCATGCAA ACCACCTGCC	7900	
CATGTGGAGC ACAGATCACC GGACATGTGA AAAACGGTTC CATGAGGATC	7950	65

# DE 199 15 178 A 1

	GTGGGGCCTA	GGACCTGTAG	TAACACGTGG	CATGGAACAT	TCCCCATTAA	8000
	CGCGTACACC	ACGGGCCCCCT	GCACGCCCTC	CCCGGCGCCA	AATTATTCTA	8050
5	GGGCGCTGTG	GCGGGTGGCT	GCTGAGGAGT	ACGTGGAGGT	TACGCGGGTG	8100
	GGGGATTTCC	ACTACGTGAC	GGGCATGACC	ACTGACAACG	TAAAGTGCCC	8150
10	GTGTCAGGTT	CCGGCCCCCG	AATCTTTCAC	AGAAGTGGAT	GGGGTGCGGT	8200
	TGCACAGGTA	CGCTCCAGCG	TGCAAACCCC	TCCTACGGGA	GGAGGTCACA	8250
15	TTCTTGGTCG	GGCTCAATCA	ATACCTGGTT	GGGTCACAGC	TCCCATGCGA	8300
	GCCCGAACCG	GACGTAGCAG	TGCTCACTTC	CATGCTCACC	GACCCCTCCC	8350
20	ACATTACGGC	GGAGACGGCT	AAGCGTAGGC	TGGCCAGGGG	ATCTCCCCCC	8400
	TCCTTGGCCA	GCTCATCAGC	TAGCCAGCTG	TCTGCGCCTT	CCTTGAAGGC	8450
	AACATGCACT	ACCCGTCATG	ACTCCCCGGA	CGCTGACCTC	ATCGAGGCCA	8500
25	ACCTCCTGTG	GCGGCAGGAG	ATGGGCGGGA	ACATCACCCG	CGTGGAGTCA	8550
	GAAAATAAGG	TAGTAATTTT	GGACTCTTTC	GAGCCGCTCC	AAGCGGAGGA	8600
30	GGATGAGAGG	GAAGTATCCG	TTCCGGCGGA	GATCCTGCGG	AGGTCCAGGA	8650
	AATTCCCTCG	AGCGATGCCC	ATATGGGCAC	GCCCGGATTA	CAACCCTCCA	8700
35	CTGTTAGAGT	CCTGGAAGGA	CCCGGACTAC	GTCCCTCCAG	TGGTACACGG	8750
	GTGTCCATTG	CCGCCTGCCA	AGGCCCTCC	GATACCACCT	CCACGGAGGA	8800
40	AGAGGACGGT	TGTCCTGTCA	GAATCTACCG	TGTCTTCTGC	CTTGGCGGAG	8850
	CTCGCCACAA	AGACCTTCGG	CAGCTCCGAA	TCGTGCGCCG	TCGACAGCGG	8900
	CACGGCAACG	GCCTCTCCTG	ACCAGCCCTC	CGACGACGGC	GACGCGGGAT	8950
45	CCGACGTTGA	GTCGTACTCC	TCCATGCCCC	CCCTTGAGGG	GGAGCCGGGG	9000
	GATCCCGATC	TCAGCGACGG	GTCTTGGTCT	ACCGTAAGCG	AGGAGGCTAG	9050
50	TGAGGACGTC	GTCTGCTGCT	CGATGTCCTA	CACATGGACA	GGCGCCCTGA	9100
	TCACGCCATG	CGCTGCGGAG	GAAACCAAGC	TGCCCATCAA	TGCACTGAGC	9150
55	AACTCTTTGC	TCCGTCACCA	CAACTTGGTC	TATGCTACAA	CATCTCGCAG	9200
	CGCAAGCCTG	CGGCAGAAGA	AGGTCACCTT	TGACAGACTG	CAGGTCCTGG	9250
	ACGACCACTA	CCGGGACGTG	CTCAAGGAGA	TGAAGGCGAA	GGCGTCCACA	9300
60	GTTAAGGCTA	AACTTCTATC	CGTGGAGGAA	GCCTGTAAGC	TGACGCCCCC	9350
	ACATTGCGCC	AGATCTAAAT	TTGGCTATGG	GGCAAAGGAC	GTCCGGAACC	9400
65	TATCCAGCAA	GGCCGTTAAC	CACATCCGCT	CCGTGTGGAA	GGACTTGCTG	9450

# DE 199 15 178 A 1

GAAGACACTG AGACACCAAT TGACACCACC ATCATGGCAA AAAATGAGGT	9500	
TTTCTGCGTC CAACCAGAGA AGGGGGGCCG CAAGCCAGCT CGCCTTATCG	9550	
TATTCCCAGA TTTGGGGGTT CGTGTGTGCG AGAAAATGGC CCTTTACGAT	9600	5
GTGGTCTCCA CCCTCCCTCA GGCCGTGATG GGCTCTTCAT ACGGATTCCA	9650	
ATACTCTCCT GGACAGCGGG TCGAGTTCCT GGTGAATGCC TGGAAAGCGA	9700	10
AGAAATGCCC TATGGGCTTC GCATATGACA CCCGCTGTTT TGACTIONACG	9750	
GTCACAGAGA ATGACATCCG TGTGAGGAG TCAATCTACC AATGTTGTGA	9800	15
CTTGGCCCCC GAAGCCAGAC AGGCCATAAG GTCGCTCACA GAGCGGCTTT	9850	
ACATCGGGGG CCCCCTGACT AATTCTAAAG GGCAGAACTG CGGCTATCGC	9900	20
CGGTGCCGCG CGAGCGGTGT ACTGACGACC AGCTGCGGTA ATACCCTCAC	9950	
ATGTTACTTG AAGGCCGCTG CGGCCTGTCG AGCTGCGAAG CTCCAGGACT	10000	
GCACGATGCT CGTATGCGGA GACGACCTTG TCGTTATCTG TGAAAGCGCG	10050	25
GGGACCCAAG AGGACGAGGC GAGCCTACGG GCCTTCACGG AGGCTATGAC	10100	
TAGATACTCT GCCCCCCTG GGGACCCGCC CAAACCAGAA TACGACTTGG	10150	30
AGTTGATAAC ATCATGCTCC TCCAATGTGT CAGTCGCGCA CGATGCATCT	10200	
GGCAAAAGGG TGTACTATCT CACCCGTGAC CCCACCACCC CCCTTGCGCG	10250	35
GGCTGCGTGG GAGACAGCTA GACACACTCC AGTCAATTCC TGGCTAGGCA	10300	
ACATCATCAT GTATGCGCCC ACCTTGTTGG CAAGGATGAT CCTGATGACT	10350	40
CATTTCTTCT CCATCCTTCT AGCTCAGGAA CAACTTGAAA AAGCCCTAGA	10400	
TTGTCAGATC TACGGGGCCT GTTACTCCAT TGAGCCACTT GACCTACCTC	10450	
AGATCATTCA ACGACTCCAT GGCCTTAGCG CATTTTCACT CCATAGTTAC	10500	45
TCTCCAGGTG AGATCAATAG GGTGGCTTCA TGCCTCAGGA AACTTGGGGT	10550	
ACCGCCCTTG CGAGTCTGGA GACATCGGGC CAGAAGTGTC CGCGCTAGGC	10600	50
TACTGTCCCA GGGGGGGAGG GCTGCCACTT GTGGCAAGTA CCTCTTCAAC	10650	
TGGGCAGTAA GGACCAAGCT CAAACTCACT CCAATCCCGG CTGCGTCCCA	10700	55
GTTGGATTTA TCCAGCTGGT TCGTTGCTGG TTACAGCGGG GGAGACATAT	10750	
ATCACAGCCT GTCTCGTGCC CGACCCCGCT GGTTCATGTG GTGCCTACTC	10800	
CTACTTTCTG TAGGGGTAGG CATCTATCTA CTCCCCAACC GATGAACGGG	10850	60
GAGCTAAACA CTCCAGGCCA ATAGGCCATC CTGTTTTTTT CCCTTTTTTT	10900	
TTTTCTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TCCTTTTTTT	10950	65

# DE 199 15 178 A 1

TTCTCTTTT TTCCTTTTC TTCCTTTGG TGGCTCCATC TTAGCCCTAG 11000

TCACGGCTAG CTGTGAAAGG TCCGTGAGCC GCTTGACTGC AGAGAGTGCT 11050

5 GATACTGGCC TCTCTGCAGA TCAAGT

10

15

20

25

30

35

40

45

50

55

60

65

# DE 199 15 178 A 1

## ANGABEN ZU SEQ ID-NO:2 :

### SEQUENZCHARAKTERISTIKA:

LÄNGE: 8637 Basenpaare

ART: Nukleinsäure

STRANGFORM: Einzelstrang

TOPOLOGIE: linear

ART DES MOLEKÜLS: Subgenom-RNA

HYPOTHETISCH: Nein

ANTI-SENSE: nein

### URSPRÜNGLICHE HERKUNFT:

ORGANISMUS: Hepatitis C Virus

STAMM: Genotyp 1b

ENTWICKLUNGSSTADIUM: —

ZELLTYP: —

### UNMITTELBARE HERKUNFT:

Isolat aus humanen Leberzellen

### MERKMAL:

NAME/SCHLÜSSEL: kodierende Sequenz für Virusproteine

LAGE: —

ERMITTLUNGSMETHODE: . experimentell

## SEQUENZBESCHREIBUNG: SEQ ID-NO: 2

I<sub>1-377</sub>/NS2-3'/wt: 8637 bp

GCCAGCCCC GATTGGGGG GACACTCCAC CATAGATCAC TCCCCTGTGA 50

GGAAGTACTG TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG 100

TGTCGTGCAG CCTCCAGGAC CCCCCCTCCC GGGAGAGCCA TAGTGGTCTG 150

CGGAACCGGT GAGTACACCG GAATTGCCAG GACGACGGG TCCTTTCTTG 200

GATCAACCCG CTCAATGCCT GGAGATTTGG GCGTGCCCC GCGAGACTGC 250

TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG 300

GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCACG 350

AATCCTAAAC CTCAAAGAAA AACCAAAGGG CGGCCATGA TTGAACAAGA 400

TGGATTGCAC GCAGGTTCTC CGGCCGCTTG GGTGGAGAGG CTATTCGGCT 450

# DE 199 15 178 A 1

	ATGACTGGGC ACAACAGACA ATCGGCTGCT CTGATGCCGC CGTGTTCGG	500
	CTGTCAGCGC AGGGGCGCCC GGTTCCTTTT GTCAAGACCG ACCTGTCCGG	550
5	TGCCCTGAAT GAACTGCAGG ACGAGGCAGC GCGGCTATCG TGGCTGGCCA	600
	CGACGGGCGT TCCTTGCGCA GCTGTGCTCG ACGTTGTCAC TGAAGCGGGA	650
10	AGGGACTGGC TGCTATTGGG CGAAGTGCCG GGGCAGGATC TCCTGTCATC	700
	TCACCTTGCT CCTGCCGAGA AAGTATCCAT CATGGCTGAT GCAATGCGGC	750
15	GGCTGCATAC GCTTGATCCG GCTACCTGCC CATTGACCA CCAAGCGAAA	800
	CATCGCATCG AGCGAGCACG TACTCGGATG GAAGCCGGTC TTGTCGATCA	850
20	GGATGATCTG GACGAAGAGC ATCAGGGGCT CGCGCCAGCC GAACTGTTTCG	900
	CCAGGCTCAA GGC GCGCATG CCCGACGGCG AGGATCTCGT CGTGACCCAT	950
25	GGCGATGCCT GCTTGCCGAA TATCATGGTG GAAAATGGCC GCTTTTCTGG	1000
	ATTCATCGAC TGTGGCCGGC TGGGTGTGGC GGACCGCTAT CAGGACATAG	1050
	CGTTGGCTAC CCGTGATATT GCTGAAGAGC TTGGCGGCGA ATGGGCTGAC	1100
30	CGCTTCCTCG TGCTTTACGG TATCGCCGCT CCCGATTCGC AGCGCATCGC	1150
	CTTCTATCGC CTTCTTGACG AGTTCTTCTG AGTTTAAACA GACCACAACG	1200
35	GTTTCCCTCT AGCGGGATCA ATTCCGCCCC TCTCCCTCCC CCCCCCTAA	1250
	CGTTACTGGC CGAAGCCGCT TGGAATAAGG CCGGTGTGCG TTTGTCTATA	1300
40	TGTTATTTTC CACCATATTG CCGTCTTTTG GCAATGTGAG GGCCCGGAAA	1350
	CCTGGCCCTG TCTTCTTGAC GAGCATTCCT AGGGGTCTTT CCCCTCTCGC	1400
45	CAAAGGAATG CAAGGTCTGT TGAATGTCGT GAAGGAAGCA GTTCCTCTGG	1450
	AAGCTTCTTG AAGACAAACA ACGTCTGTAG CGACCCTTTG CAGGCAGCGG	1500
	AACCCCCAC CTGGCGACAG GTGCCTCTGC GGCCAAAAGC CACGTGTATA	1550
50	AGATACACCT GCAAAGGCGG CACAACCCCA GTGCCACGTT GTGAGTTGGA	1600
	TAGTTGTGGA AAGAGTCAA TGGCTCTCCT CAAGCGTATT CAACAAGGGG	1650
55	CTGAAGGATG CCCAGAAGGT ACCCCATTGT ATGGGATCTG ATCTGGGGCC	1700
	TCGGTGCACA TGCTTTACAT GTGTTTAGTC GAGGTAAAA AACGTCTAGG	1750
60	CCCCCGAAC CACGGGGACG TGGTTTTCCT TTGAAAAACA CGATAATACC	1800
	ATGGACCGGG AGATGGCAGC ATCGTGCGGA GGCGCGGTTT TCGTAGGTCT	1850
	GATACTCTTG ACCTTGTCAC CGCACTATAA GCTGTTCCCTC GCTAGGCTCA	1900
65	TATGGTGGTT ACAATATTTT ATCACCAGGG CCGAGGCACA CTTGCAAGTG	1950



# DE 199 15 178 A 1

TGGATCCCC	CCCTCAACGT	TCGGGGGGGC	CGCGATGCCG	TCATCCTCCT	2000	
CACGTGCGCG	ATCCACCCAG	AGCTAATCTT	TACCATCACC	AAAATCTTGC	2050	
TCGCCATACT	CGGTCCACTC	ATGGTGCTCC	AGGCTGGTAT	AACCAAAGTG	2100	5
CCGTACTTCG	TGCGCGCACA	CGGGCTCATT	CGTGCGATGCA	TGCTGGTGCG	2150	
GAAGGTTGCT	GGGGGTCATT	ATGTCCAAAT	GGCTCTCATG	AAGTTGGCCG	2200	10
CACTGACAGG	TACGTACGTT	TATGACCATC	TCACCCCACT	GCGGGACTGG	2250	
GCCCACGCGG	GCCTACGAGA	CCTTGCGGTG	GCAGTTGAGC	CCGTCGTCTT	2300	15
CTCTGATATG	GAGACCAAGG	TTATCACCTG	GGGGGCAGAC	ACCGCGGCGT	2350	
GTGGGGACAT	CATCTTGGGC	CTGCCCCTCT	CCGCCCCGAG	GGGGAGGGAG	2400	20
ATACATCTGG	GACCGGCAGA	CAGCCTTGAA	GGGCAGGGGT	GGCGACTCCT	2450	
CGCGCCTATT	ACGGCCTACT	CCCAACAGAC	GCGAGGCCTA	CTTGGCTGCA	2500	25
TCATCACTAG	CCTCACAGGC	CGGGACAGGA	ACCAGGTCGA	GGGGGAGGTC	2550	
CAAGTGGTCT	CCACCGCAAC	ACAATCTTTC	CTGGCGACCT	GCGTCAATGG	2600	
CGTGTGTTGG	ACTGTCTATC	ATGGTGCCGG	CTCAAAGACC	CTTGCCGGCC	2650	30
CAAAGGGCCC	AATCACCCAA	ATGTACACCA	ATGTGGACCA	GGACCTCGTC	2700	
GGCTGGCAAG	CGCCCCCGG	GGCGCGTTCC	TTGACACCAT	GCACCTGGGG	2750	35
CAGCTCGGAC	CTTTACTTGG	TCACGAGGCA	TGCCGATGTC	ATTCCGGTGC	2800	
GCCGGCGGGG	CGACAGCAGG	GGGAGCCTAC	TCTCCCCCAG	GCCCCTCTCC	2850	40
TACTTGAAGG	GCTCTTCGGG	CGGTCCACTG	CTCTGCCCCCT	CGGGGCACGC	2900	
TGTGGGCATC	TTTCGGGCTG	CCGTGTGCAC	CCGAGGGGTT	GCGAAGGCGG	2950	
TGGACTTTGT	ACCCGTCGAG	TCTATGGAAA	CCACTATGCG	GTCCCCGGTC	3000	45
TTCACGGACA	ACTCGTCCCC	TCCGGCCGTA	CCGCAGACAT	TCCAGGTGGC	3050	
CCATCTACAC	GCCCCTACTG	GTAGCGGCAA	GAGCACTAAG	GTGCCGGCTG	3100	50
CGTATGCAGC	CCAAGGGTAT	AAGGTGCTTG	TCCTGAACCC	GTCCGTCGCC	3150	
GCCACCCTAG	GTTTCGGGGC	GTATATGTCT	AAGGCACATG	GTATCGACCC	3200	55
TAACATCAGA	ACCGGGGTAA	GGACCATCAC	CACGGGTGCC	CCCATCACGT	3250	
ACTCCACCTA	TGGCAAGTTT	CTTGCCGACG	GTGGTTGCTC	TGGGGGCGCC	3300	60
TATGACATCA	TAATATGTGA	TGAGTGCCAC	TCAACTGACT	CGACCACTAT	3350	
CCTGGGCATC	GGCACAGTCC	TGGACCAAGC	GGAGACGGCT	GGAGCGCGAC	3400	
TCGTCGTGCT	CGCCACCGCT	ACGCCTCCGG	GATCGGTCAC	CGTGCCACAT	3450	65

# DE 199 15 178 A 1

	CCAAACATCG	AGGAGGTGGC	TCTGTCCAGC	ACTGGAGAAA	TCCCCTTTTA	3500
	TGGCAAAGCC	ATCCCCATCG	AGACCATCAA	GGGGGGGAGG	CACCTCATTT	3550
5	TCTGCCATTC	CAAGAAGAAA	TGTGATGAGC	TCGCCGCGAA	GCTGTCCGGC	3600
	CTCGGACTCA	ATGCTGTAGC	ATATTACCGG	GGCCTTGATG	TATCCGTCAT	3650
10	ACCAACTAGC	GGAGACGTCA	TTGTGCTAGC	AACGGACGCT	CTAATGACGG	3700
	GCTTTACCGG	CGATTTTCGAC	TCAGTGATCG	ACTGCAATAC	ATGTGTCACC	3750
15	CAGACAGTCG	ACTTCAGCCT	GGACCCGACC	TTCACCATTG	AGACGACGAC	3800
	CGTGCCACAA	GACGCGGTGT	CACGCTCGCA	GCGGCGAGGC	AGGACTGGTA	3850
20	GGGGCAGGAT	GGGCATTTAC	AGGTTTGTGA	CTCCAGGAGA	ACGGCCCTCG	3900
	GGCATGTTCG	ATTCTTCGGT	TCTGTGCGAG	TGCTATGACG	CGGGCTGTGC	3950
	TTGGTACGAG	CTCACGCCCG	CCGAGACCTC	AGTTAGGTTG	CGGGCTTACC	4000
25	TAAACACACC	AGGGTTGCCC	GTCTGCCAGG	ACCATCTGGA	GTTCTGGGAG	4050
	AGCGTCTTTA	CAGGCCTCAC	CCACATAGAC	GCCCATTCTT	TGTCCCAGAC	4100
30	TAAGCAGGCA	GGAGACAACT	TCCCCTACCT	GGTAGCATAC	CAGGCTACGG	4150
	TGTGCGCCAG	GGCTCAGGCT	CCACCTCCAT	CGTGGGACCA	AATGTGGAAG	4200
35	TGTCTCATAC	GGCTAAAGCC	TACGCTGCAC	GGGCCAACGC	CCCTGCTGTA	4250
	TAGGCTGGGA	GCCGTTCAAA	ACGAGGTTAC	TACCACACAC	CCCATAACCA	4300
40	AATACATCAT	GGCATGCATG	TCGGCTGACC	TGGAGGTCGT	CACGAGCACC	4350
	TGGGTGCTGG	TAGGCGGAGT	CCTAGCAGCT	CTGGCCGCGT	ATTGCCTGAC	4400
	AACAGGCAGC	GTGGTCATTG	TGGGCAGGAT	CATCTTGTCC	GGAAAGCCGG	4450
45	CCATCATTCC	CGACAGGGAA	GTCCTTTACC	GGGAGTTCGA	TGAGATGGAA	4500
	GAGTGCGCCT	CACACCTCCC	TTACATCGAA	CAGGGAATGC	AGCTCGCCGA	4550
50	ACAATTCAAA	CAGAAGGCAA	TCGGGTTGCT	GCAAACAGCC	ACCAAGCAAG	4600
	CGGAGGCTGC	TGCTCCCGTG	GTGGAATCCA	AGTGGCGGAC	CCTCGAAGCC	4650
55	TTCTGGGCGA	AGCATATGTG	GAATTTCATC	AGCGGGATAC	AATATTTAGC	4700
	AGGCTTGCTC	ACTCTGCCTG	GCAACCCCGC	GATAGCATCA	CTGATGGCAT	4750
	TCACAGCCTC	TATCACCAGC	CCGCTCACCA	CCCAACATAC	CCTCCTGTTT	4800
60	AACATCCTGG	GGGGATGGGT	GGCCGCCCAA	CTTGCTCCTC	CCAGCGCTGC	4850
	TTCTGCTTTC	GTAGGCGCCG	GCATCGCTGG	AGCGGCTGTT	GGCAGCATAG	4900
65	GCCTTGGGAA	GGTGCTTGTC	GATATTTTGG	CAGGTTATGG	AGCAGGGGTG	4950

# DE 199 15 178 A 1

GCAGGCGCGC TCGTGGCCTT TAAGGTCATG AGCGGCGAGA TGCCCTCCAC	5000	
CGAGGACCTG GTTAACCTAC TCCCTGCTAT CCTCTCCCCT GCGGCCCTAG	5050	
TCGTGCGGGT CGTGTGCGCA GCGATACTGC GTCGGCACGT GGGCCCAGGG	5100	5
GAGGGGGCTG TGCAGTGGAT GAACCGGCTG ATAGCGTTCG CTTCGCGGGG	5150	
TAACCACGTC TCCCCACGC ACTATGTGCC TGAGAGCGAC GCTGCAGCAC	5200	10
GTGTCACTCA GATCCTCTCT AGTCTTACCA TCACTCAGCT GCTGAAGAGG	5250	
CTTCACCAGT GGATCAACGA GGA CTGCTCC ACGCCATGCT CCGGCTCGTG	5300	15
GCTAAGAGAT GTTTGGGATT GGATATGCAC GGTGTTGACT GATTTCAGA	5350	
CCTGGCTCCA GTCCAAGCTC CTGCCGCGAT TGCCGGGAGT CCCCTTCTTC	5400	20
TCATGTCAAC GTGGGTACAA GGGAGTCTGG CGGGGCGACG GCATCATGCA	5450	
AACCACCTGC CCATGTGGAG CACAGATCAC CGGACATGTG AAAAACGGTT	5500	
CCATGAGGAT CGTGGGGCCT AGGACCTGTA GTAACACGTG GCATGGAACA	5550	25
TTCCCCATTA ACGCGTACAC CACGGGCCCC TGACGCCCCC CCCC GGCGCC	5600	
AAATTATTCT AGGGCGCTGT GGGGGGTGGC TGCTGAGGAG TACGTGGAGG	5650	30
TTACGCGGGT GGGGGATTTC CACTACGTGA CGGGCATGAC CACTGACAAC	5700	
GTAAAGTGCC CGTGTGAGGT TCCGGCCCCC GAATTCTTCA CAGAAGTGGA	5750	35
TGGGGTGCGG TTGCACAGGT ACGCTCCAGC GTGCAAACCC CTCCTACGGG	5800	
AGGAGGTCAC ATTCTGGTC GGGCTCAATC AATACCTGGT TGGGTCACAG	5850	40
CTCCCATGCG AGCCCGAACC GGACGTAGCA GTGCTCACTT CCATGCTCAC	5900	
CGACCCCTCC CACATTACGG CGGAGACGGC TAAGCGTAGG CTGGCCAGGG	5950	
GATCTCCCCC CTCCTTGGCC AGCTCATCAG CTAGCCAGCT GTCTGCGCCT	6000	45
TCCTTGAAGG CAACATGCAC TACCCGTCAT GACTCCCCGG ACGCTGACCT	6050	
CATCGAGGCC AACCTCCTGT GCGGCGAGGA GATGGGCGGG AACATCACCC	6100	50
GCGTGGAGTC AGAAAATAAG GTAGTAATTT TGGACTCTTT CGAGCCGCTC	6150	
CAAGCGGAGG AGGATGAGAG GGAAGTATCC GTTCCGGCGG AGATCCTGCG	6200	55
GAGGTCCAGG AAATTCCCTC GAGCGATGCC CATATGGGCA CGCCCGGATT	6250	
ACAACCCTCC ACTGTTAGAG TCCTGGAAGG ACCCGGACTA CGTCCCTCCA	6300	
GTGGTACACG GGTGTCCATT GCCGCCTGCC AAGGCCCTC CGATACCACC	6350	60
TCCACGGAGG AAGAGGACGG TTGTCCTGTC AGAATCTACC GTGTCTTCTG	6400	
CCTTGGCGGA GCTCGCCACA AAGACCTTCG GCAGCTCCGA ATCGTCGGCC	6450	65

# DE 199 15 178 A 1

	GTCGACAGCG	GCACGGCAAC	GGCCTCTCCT	GACCAGCCCT	CCGACGACGG	6500
	CGACGCGGGA	TCCGACGTTG	AGTCGTACTC	CTCCATGCCC	CCCCTTGAGG	6550
5	GGGAGCCGGG	GGATCCCGAT	CTCAGCGACG	GGTCTTGGTC	TACCGTAAGC	6600
	GAGGAGGCTA	GTGAGGACGT	CGTCTGCTGC	TCGATGTCCT	ACACATGGAC	6650
10	AGGCGCCCTG	ATCACGCCAT	GCGCTGCGGA	GGAAACCAAG	CTGCCCATCA	6700
	ATGCACTGAG	CAACTCTTTG	CTCCGTCACC	ACAACTTGGT	CTATGCTACA	6750
15	ACATCTCGCA	GCGCAAGCCT	GCGGCAGAAG	AAGGTCACCT	TTGACAGACT	6800
	GCAGGTCCTG	GACGACCACT	ACCGGGACGT	GCTCAAGGAG	ATGAAGGCGA	6850
20	AGGCGTCCAC	AGTTAAGGCT	AAACTTCTAT	CCGTGGAGGA	AGCCTGTAAG	6900
	CTGACGCCCC	CACATTCGGC	CAGATCTAAA	TTTGGCTATG	GGGCAAAGGA	6950
	CGTCCGGAAC	CTATCCAGCA	AGGCCGTTAA	CCACATCCGC	TCCGTGTGGA	7000
25	AGGACTTGCT	GGAAGACACT	GAGACACCAA	TTGACACCAC	CATCATGGCA	7050
	AAAAATGAGG	TTTTCTGCGT	CCAACCAGAG	AAGGGGGGCC	GCAAGCCAGC	7100
30	TCGCCTTATC	GTATCCCGAG	ATTTGGGGGT	TCGTGTGTGC	GAGAAAATGG	7150
	CCCTTTACGA	TGTGGTCTCC	ACCCTCCCTC	AGGCCGTGAT	GGGCTCTTCA	7200
35	TACGGATTCC	AATACTCTCC	TGGACAGCGG	GTCGAGTTCC	TGGTGAATGC	7250
	CTGGAAAGCG	AAGAAATGCC	CTATGGGCTT	CGCATATGAC	ACCCGCTGTT	7300
40	TTGACTCAAC	GGTCACTGAG	AATGACATCC	GTGTTGAGGA	GTCAATCTAC	7350
	CAATGTTGTG	ACTTGGCCCC	CGAAGCCAGA	CAGGCCATAA	GGTCGCTCAC	7400
	AGAGCGGCTT	TACATCGGGG	GCCCCCTGAC	TAATTCTAAA	GGGCAGAACT	7450
45	GCGGCTATCG	CCGGTGCCGC	GCGAGCGGTG	TACTGACGAC	CAGCTGCGGT	7500
	AATACCCTCA	CATGTTACTT	GAAGGCCGCT	GCGGCCTGTC	GAGCTGCGAA	7550
50	GCTCCAGGAC	TGCACGATGC	TCGTATGCGG	AGACGACCTT	GTCGTTATCT	7600
	GTGAAAGCGC	GGGGACCCAA	GAGGACGAGG	CGAGCCTACG	GGCCTTCACG	7650
55	GAGGCTATGA	CTAGATACTC	TGCCCCCCTT	GGGGACCCGC	CCAAACCAGA	7700
	ATACGACTTG	GAGTTGATAA	CATCATGCTC	CTCCAATGTG	TCAGTCGCGC	7750
	ACGATGCATC	TGGCAAAAGG	GTGTACTATC	TCACCCGTGA	CCCCACCACC	7800
60	CCCCTTGCGC	GGGCTGCGTG	GGAGACAGCT	AGACACACTC	CAGTCAATTC	7850
	CTGGCTAGGC	AACATCATCA	TGTATGCGCC	CACCTTGTGG	GCAAGGATGA	7900
65	TCCTGATGAC	TCATTTCTTC	TCCATCCTTC	TAGCTCAGGA	ACAACTTGAA	7950

# DE 199 15 178 A 1

AAAGCCCTAG ATTGTCAGAT CTACGGGGCC TGTACTCCA TTGAGCCACT	8000	
TGACCTACCT CAGATCATT C AACGACTCCA TGGCCTTAGC GCATTTTCAC	8050	
TCCATAGTTA CTCTCCAGGT GAGATCAATA GGGTGGCTTC ATGCCTCAGG	8100	5
AAACTTGGGG TACCGCCCTT GCGAGTCTGG AGACATCGGG CCAGAAGTGT	8150	
CCGCGCTAGG CTACTGTCC AGGGGGGGAG GGCTGCCACT TGTGGCAAGT	8200	10
ACCTCTTCAA CTGGGCAGTA AGGACCAAGC TCAAACCTCAC TCCAATCCCG	8250	
GCTGCGTCCC AGTTGGATTT ATCCAGCTGG TTCGTTGCTG GTTACAGCGG	8300	15
GGGAGACATA TATCACAGCC TGTCTCGTGC CCGACCCCGC TGGTTCATGT	8350	
GGTGCCTACT CCTACTTTCT GTAGGGGTAG GCATCTATCT ACTCCCCAAC	8400	20
CGATGAACGG GGAGCTAAAC ACTCCAGGCC AATAGGCCAT CCTGTTTTTT	8450	
TCCCTTTTTT TTTTCTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT	8500	25
CTCCTTTTTT TTTCTCTTT TTTTCTTTT CTTTCTTTG GTGGCTCCAT	8550	
CTTAGCCCTA GTCACGGCTA GCTGTGAAAG GTCCGTGAGC CGCTTGACTG	8600	
CAGAGAGTGC TGATACTGGC CTCTCTGCAG ATCAAGT		30
		35
		40
		45
		50
		55
		60
		65

ANGABEN ZU SEQ ID-NO:3 :

5 SEQUENZCHARAKTERISTIKA:  
 LÄNGE: 8001 Basenpaare  
 ART: Nukleinsäure  
 STRANGFORM: Einzelstrang  
 10 TOPOLOGIE: linear

ART DES MOLEKÜLS: Subgenom-RNA

15 HYPOTHETISCH: Nein

ANTI-SENSE: nein

20 URSPRÜNGLICHE HERKUNFT:  
 ORGANISMUS: Hepatitis C Virus  
 STAMM: Genotyp 1b  
 25 ENTWICKLUNGSSTADIUM: —  
 ZELLTYP: —

UNMITTELBARE HERKUNFT:  
 30 Isolat aus humanen Leberzellen

MERKMAL:  
 NAME/SCHLÜSSEL: kodierende Sequenz für Virusproteine  
 35 LAGE: —  
 ERMITTLUNGSMETHODE: . experimentell

40

45 SEQUENZBESCHREIBUNG: SEQ ID-NO: 3

50 I<sub>389</sub>/NS3-3' /wt: 8001 bp

GCCAGCCCC	GATTGGGGG	GACACTCCAC	CATAGATCAC	TCCCCTGTGA	50
GGAAC TACTG	TCTTCACGA	GAAAGCGTCT	AGCCATGGCG	TTAGTATGAG	100
55 TGTCGTGCAG	CCTCCAGGAC	CCCCCCTCCC	GGGAGAGCCA	TAGTGGTCTG	150
CGGAACCGGT	GAGTACACCG	GAATTGCCAG	GACGACCGGG	TCCTTTCTTG	200
60 GATCAACCCG	CTCAATGCCT	GGAGATT TGG	GCGTGCCCCC	GCGAGACTGC	250
TAGCCGAGTA	GTGTTGGGTC	GCGAAAGGCC	TTGTGGTACT	GCCTGATAGG	300
65 GTGCTTGCGA	GTGCCCCGGG	AGGTCTCGTA	GACCGTGCAC	CATGAGCACG	350
AATCCTAAAC	CTCAAAGAAA	AACCAAACGT	AACACCAACG	GGCGCGCCAT	400

# DE 199 15 178 A 1

GATTGAACAA	GATGGATTGC	ACGCAGGTTT	TCCGGCCGCT	TGGGTGGAGA	450	
GGCTATTTCG	CTATGACTGG	GCACAACAGA	CAATCGGCTG	CTCTGATGCC	500	
GCCGTGTTCC	GGCTGTCAGC	GCAGGGGCGC	CCGGTTCTTT	TTGTCAAGAC	550	5
CGACCTGTCC	GGTGCCCTGA	ATGAACTGCA	GGACGAGGCA	GCGCGGCTAT	600	
CGTGGCTGGC	CACGACGGGC	GTTCCCTTGC	CAGCTGTGCT	CGACGTTGTC	650	10
ACTGAAGCGG	GAAGGGACTG	GCTGCTATTG	GGCGAAGTGC	CGGGGCAGGA	700	
TCTCCTGTCA	TCTCACCTTG	CTCCTGCCGA	GAAAGTATCC	ATCATGGCTG	750	15
ATGCAATGCG	GCGGCTGCAT	ACGCTTGATC	CGGCTACCTG	CCCATTCGAC	800	
CACCAAGCGA	AACATCGCAT	CGAGCGAGCA	CGTACTCGGA	TGGAAGCCGG	850	20
TCTTGTCGAT	CAGGATGATC	TGGACGAAGA	GCATCAGGGG	CTCGCGCCAG	900	
CCGAACTGTT	CGCCAGGCTC	AAGGCGCGCA	TGCCCAGACG	CGAGGATCTC	950	
GTCGTGACCC	ATGGCGATGC	CTGCTTGCCG	AATATCATGG	TGGAATATGG	1000	25
CCGCTTTTCT	GGATTCATCG	ACTGTGGCCG	GCTGGGTGTG	GCGGACCGCT	1050	
ATCAGGACAT	AGCGTTGGCT	ACCCGTGATA	TTGCTGAAGA	GCTTGGCGGC	1100	30
GAATGGGCTG	ACCGCTTCCT	CGTGCTTTAC	GGTATCGCCG	CTCCCGATTC	1150	
GCAGCGCATC	GCCTTCTATC	GCCTTCTTGA	CGAGTTCTTC	TGAGTTTAAA	1200	35
CAGACCACAA	CGGTTTCCCT	CTAGCGGGAT	CAATTCCGCC	CCTCTCCCTC	1250	
CCCCCCCCCT	AACGTTACTG	GCCGAAGCCG	CTTGAATAA	GGCCGGTGTG	1300	40
CGTTTGCTA	TATGTTATTT	TCCACCATAT	TGCCGTCTTT	TGGCAATGTG	1350	
AGGGCCCGGA	AACCTGGCCC	TGTCTTCTTG	ACGAGCATTC	CTAGGGGTCT	1400	
TTCCCTCTC	GCCAAAGGAA	TGCAAGGTCT	GTTGAATGTC	GTGAAGGAAG	1450	45
CAGTTCCTCT	GGAAGCTTCT	TGAAGACAAA	CAACGTCTGT	AGCGACCCTT	1500	
TGCAGGCAGC	GGAACCCCCC	ACCTGGCGAC	AGGTGCCTCT	GCGGCCAAAA	1550	50
GCCACGTGTA	TAAGATACAC	CTGCAAAGGC	GGCACAACCC	CAGTGCCACG	1600	
TTGTGAGTTG	GATAGTTGTG	GAAAGAGTCA	AATGGCTCTC	CTCAAGCGTA	1650	55
TTCAACAAGG	GGCTGAAGGA	TGCCCAGAAG	GTACCCCAT	GTATGGGATC	1700	
TGATCTGGGG	CCTCGGTGCA	CATGCTTTAC	ATGTGTTTAG	TCGAGGTTAA	1750	
AAAACGTCTA	GGCCCCCGA	ACCACGGGGA	CGTGGTTTTT	CTTTGAAAAA	1800	60
CACGATAATA	CCATGGCGCC	TATTACGGCC	TACTCCCAAC	AGACGCGAGG	1850	
CCTACTTGGC	TGCATCATCA	CTAGCCTCAC	AGGCCGGGAC	AGGAACCAGG	1900	65

# DE 199 15 178 A 1

	TCGAGGGGGA	GGTCCAAGTG	GTCTCCACCG	CAACACAATC	TTTCCTGGCG	1950
	ACCTGCGTCA	ATGGCGTGTG	TTGGACTGTC	TATCATGGTG	CCGGCTCAAA	2000
5	GACCCTTGCC	GGCCCAAAGG	GCCCAATCAC	CCAAATGTAC	ACCAATGTGG	2050
	ACCAGGACCT	CGTCGGCTGG	CAAGCGCCCC	CCGGGGCGCG	TTCTTGACA	2100
10	CCATGCACCT	GCGGCAGCTC	GGACCTTTAC	TTGGTCACGA	GGCATGCCGA	2150
	TGTCATTCCG	GTGCGCCGGC	GGGGCGACAG	CAGGGGGAGC	CTACTCTCCC	2200
15	CCAGGCCCGT	CTCCTACTTG	AAGGGCTCTT	CGGGCGGTCC	ACTGCTCTGC	2250
	CCCTCGGGGC	ACGCTGTGGG	CATCTTTCGG	GCTGCCGTGT	GCACCCGAGG	2300
20	GGTTGCGAAG	GCGGTGGACT	TTGTACCCGT	CGAGTCTATG	GAAACCACTA	2350
	TGCGGTCCCC	GGTCTTCACG	GACAACTCGT	CCCCTCCGGC	CGTACCGCAG	2400
	ACATTCCAGG	TGGCCCATCT	ACACGCCCCCT	ACTGGTAGCG	GCAAGAGCAC	2450
25	TAAGGTGCCG	GCTGCGTATG	CAGCCCAAGG	GTATAAGGTG	CTTGTCCTGA	2500
	ACCCGTCCGT	CGCCGCCACC	CTAGGTTTCG	GGGCGTATAT	GTCTAAGGCA	2550
30	CATGGTATCG	ACCCTAACAT	CAGAACCGGG	GTAAGGACCA	TCACCACGGG	2600
	TGCCCCCATC	ACGTACTCCA	CCTATGGCAA	GTTTCTTGCC	GACGGTGGTT	2650
35	GCTCTGGGGG	CGCCTATGAC	ATCATAATAT	GTGATGAGTG	CCACTCAACT	2700
	GACTCGACCA	CTATCCTGGG	CATCGGCACA	GTCCTGGACC	AAGCGGAGAC	2750
40	GGCTGGAGCG	CGACTCGTCG	TGCTCGCCAC	CGCTACGCCT	CCGGGATCGG	2800
	TCACCGTGCC	ACATCCAAAC	ATCGAGGAGG	TGGCTCTGTC	CAGCACTGGA	2850
	GAAATCCCCT	TTTATGGCAA	AGCCATCCCC	ATCGAGACCA	TCAAGGGGGG	2900
45	GAGGCACCTC	ATTTTCTGCC	ATTCCAAGAA	GAAATGTGAT	GAGCTCGCCG	2950
	CGAAGCTGTC	CGGCCTCGGA	CTCAATGCTG	TAGCATATTA	CCGGGGCCTT	3000
50	GATGTATCCG	TCATACCAAC	TAGCGGAGAC	GTCATTGTCT	TAGCAACGGA	3050
	CGCTCTAATG	ACGGGCTTTA	CCGGCGATTT	CGACTCAGTG	ATCGACTGCA	3100
55	ATACATGTGT	CACCCAGACA	GTCGACTTCA	GCCTGGACCC	GACCTTCACC	3150
	ATTGAGACGA	CGACCGTGCC	ACAAGACGCG	GTGTCACGCT	CGCAGCGGCG	3200
	AGGCAGGACT	GGTAGGGGCA	GGATGGGCAT	TTACAGGTTT	GTGACTCCAG	3250
60	GAGAACGGCC	CTCGGGCATG	TTCGATTCCCT	CGGTTCTGTG	CGAGTGCTAT	3300
	GACGCGGGCT	GTGCTTG GTA	CGAGCTCACG	CCCGCCGAGA	CCTCAGTTAG	3350
65	GTTGCGGGCT	TACCTAAACA	CACCAGGGTT	GCCCGTCTGC	CAGGACCATC	3400



# DE 199 15 178 A 1

TGGAGTTCTG	GGAGAGCGTC	TTTACAGGCC	TCACCCACAT	AGACGCCCAT	3450	
TTCTTGTC	AGACTAAGCA	GGCAGGAGAC	AACTTCCCCT	ACCTGGTAGC	3500	
ATACCAGGCT	ACGGTGTGCG	CCAGGGCTCA	GGCTCCACCT	CCATCGTGGG	3550	5
ACCAAATGTG	GAAGTGTCTC	ATACGGCTAA	AGCCTACGCT	GCACGGGCCA	3600	
ACGCCCCCTG	TGTATAGGCT	GGGAGCCGTT	CAAAACGAGG	TTACTACCAC	3650	10
ACACCCCAT	ACCAAATACA	TCATGGCATG	CATGTCGGCT	GACCTGGAGG	3700	
TCGTCACGAG	CACCTGGGTG	CTGGTAGGCG	GAGTCCTAGC	AGCTCTGGCC	3750	15
GCGTATTGCC	TGACAACAGG	CAGCGTGGTC	ATTGTGGGCA	GGATCATCTT	3800	
GTCCGGAAAG	CCGGCCATCA	TTCCCGACAG	GGAAGTCCTT	TACCGGGAGT	3850	20
TCGATGAGAT	GGAAGAGTGC	GCCTCACACC	TCCCTTACAT	CGAACAGGGA	3900	
ATGCAGCTCG	CCGAACAATT	CAAACAGAAG	GCAATCGGGT	TGCTGCAAAC	3950	
AGCCACCAAG	CAAGCGGAGG	CTGCTGCTCC	CGTGGTGGAA	TCCAAGTGGC	4000	25
GGACCCTCGA	AGCCTTCTGG	GCGAAGCATA	TGTGGAATTT	CATCAGCGGG	4050	
ATACAATATT	TAGCAGGCTT	GTCCACTCTG	CCTGGCAACC	CCGCGATAGC	4100	30
ATCACTGATG	GCATTACAG	CCTCTATCAC	CAGCCCCTC	ACCACCCAAC	4150	
ATACCCTCCT	GTTTAACATC	CTGGGGGGAT	GGGTGGCCGC	CCAATTGCT	4200	35
CCTCCCAGCG	CTGCTTCTGC	TTTCGTAGGC	GCCGGCATCG	CTGGAGCGGC	4250	
TGTTGGCAGC	ATAGGCCTTG	GGAAGGTGCT	TGTGGATATT	TTGGCAGGTT	4300	40
ATGGAGCAGG	GGTGGCAGGC	GCGCTCGTGG	CCTTTAAGGT	CATGAGCGGC	4350	
GAGATGCCCT	CCACCGAGGA	CCTGGTTAAC	CTACTCCCTG	CTATCCTCTC	4400	
CCCTGGCGCC	CTAGTCGTGC	GGGTCGTGTG	CGCAGCGATA	CTGCGTCGGC	4450	45
ACGTGGGCCC	AGGGGAGGGG	GCTGTGCAGT	GGATGAACCG	GCTGATAGCG	4500	
TTGCGTTTCG	GGGGTAACCA	CGTCTCCCCC	ACGCACTATG	TGCCTGAGAG	4550	50
CGACGCTGCA	GCACGTGTCA	CTCAGATCCT	CTCTAGTCTT	ACCATCACTC	4600	
AGCTGCTGAA	GAGGCTTCAC	CAGTGGATCA	ACGAGGACTG	CTCCACGCCA	4650	55
TGCTCCGGCT	CGTGGCTAAG	AGATGTTTGG	GATTGGATAT	GCACGGTGTT	4700	
GACTGATTTT	AAGACCTGGC	TCCAGTCCAA	GCTCCTGCCG	CGATTGCCGG	4750	
GAGTCCCCTT	CTTCTCATGT	CAACGTGGGT	ACAAGGGAGT	CTGGCGGGGC	4800	60
GACGGCATCA	TGCAAACCAC	CTGCCCATGT	GGAGCACAGA	TCACCGGACA	4850	
TGTGAAAAAC	GGTTCCATGA	GGATCGTGGG	GCCTAGGACC	TGTAGTAACA	4900	65

# DE 199 15 178 A 1

	CGTGGCATGG AACATTCCCC ATTAACGCGT ACACCACGGG CCCCTGCACG	4950
	CCCTCCCCGG CGCCAAATTA TTCTAGGGCG CTGTGGCGGG TGGCTGCTGA	5000
5	GGAGTACGTG GAGGTTACGC GGGTGGGGGA TTTCCACTAC GTGACGGGCA	5050
	TGACCACTGA CAACGTAAAG TGCCCGTGTC AGGTTCCGGC CCCCGAATTC	5100
10	TTCACAGAAG TGGATGGGGT GCGGTTGCAC AGGTACGCTC CAGCGTGCAA	5150
	ACCCCTCCTA CGGGAGGAGG TCACATTCCT GGTCGGGCTC AATCAATACC	5200
15	TGGTTGGGTC ACAGCTCCCA TGCGAGCCCG AACCGGACGT AGCAGTGCTC	5250
	ACTTCCATGC TCACCGACCC CTCCCACATT ACGGCGGAGA CGGCTAAGCG	5300
20	TAGGCTGGCC AGGGGATCTC CCCCTCCTT GGCCAGCTCA TCAGCTAGCC	5350
	AGCTGTCTGC GCCTTCCTTG AAGGCAACAT GCACTACCCG TCATGACTCC	5400
	CCGGACGCTG ACCTCATCGA GGCCAACCTC CTGTGGCGGC AGGAGATGGG	5450
25	CGGGAACATC ACCCGCGTGG AGTCAGAAAA TAAGGTAGTA ATTTTGGACT	5500
	CTTTCGAGCC GCTCCAAGCG GAGGAGGATG AGAGGGAAGT ATCCGTTCCG	5550
30	GCGGAGATCC TGCGGAGGTC CAGGAAATTC CCTCGAGCGA TGCCCATATG	5600
	GGCAGCCCCG GATTACAACC CTCCACTGTT AGAGTCCTGG AAGGACCCGG	5650
35	ACTACGTCCC TCCAGTGGTA CACGGGTGTC CATTGCCGCC TGCCAAGGCC	5700
	CCTCCGATAC CACCTCCACG GAGGAAGAGG ACGGTTGTCC TGTCAGAATC	5750
40	TACCGTGTCT TCTGCCTTGG CGGAGCTCGC CACAAAGACC TTCGGCAGCT	5800
	CCGAATCGTC GGCCGTCGAC AGCGGCACGG CAACGGCCTC TCCTGACCAG	5850
	CCCTCCGACG ACGGCGACGC GGGATCCGAC GTTGAGTCGT ACTCCTCCAT	5900
45	GCCCCCCTT GAGGGGGAGC CGGGGGATCC CGATCTCAGC GACGGGTCTT	5950
	GGTCTACCGT AAGCGAGGAG GCTAGTGAGG ACGTCGTCTG CTGCTCGATG	6000
50	TCCTACACAT GGACAGGCGC CCTGATCAGC CCATGCGCTG CGGAGGAAAC	6050
	CAAGCTGCCC ATCAATGCAC TGAGCAACTC TTGCTCCGT CACCACAAC	6100
55	TGGTCTATGC TACAACATCT CGCAGCGCAA GCCTGCGGCA GAAGAAGGTC	6150
	ACCTTTGACA GACTGCAGGT CCTGGACGAC CACTACCGGG ACGTGCTCAA	6200
	GGAGATGAAG GCGAAGGCGT CCACAGTTAA GGCTAAACTT CTATCCGTGG	6250
60	AGGAAGCCTG TAAGCTGACG CCCCACATT CGGCCAGATC TAAATTTGGC	6300
	TATGGGGCAA AGGACGTCCG GAACCTATCC AGCAAGGCCG TTAACCACAT	6350
65	CCGCTCCGTG TGGAAGGACT TGCTGGAAGA CACTGAGACA CCAATTGACA	6400

# DE 199 15 178 A 1

CCACCATCAT	GGCAAAAAT	GAGGTTTTCT	GCGTCCAACC	AGAGAAGGGG	6450	
GGCCGCAAGC	CAGCTCGCCT	TATCGTATTC	CCAGATTTGG	GGGTTCTGT	6500	
GTGCGAGAAA	ATGGCCCTTT	ACGATGTGGT	CTCCACCCTC	CCTCAGGCCG	6550	5
TGATGGGCTC	TTCATACGGA	TTCCAATACT	CTCCTGGACA	GCGGGTCGAG	6600	
TTCTTGGTGA	ATGCCTGGAA	AGCGAAGAAA	TGCCCTATGG	GCTTCGCATA	6650	10
TGACACCCGC	TGTTTTGACT	CAACGGTCAC	TGAGAATGAC	ATCCGTGTTG	6700	
AGGAGTCAAT	CTACCAATGT	TGTGACTTGG	CCCCCGAAGC	CAGACAGGCC	6750	15
ATAAGGTCGC	TCACAGAGCG	GCTTTACATC	GGGGGCCCCC	TGACTAATTC	6800	
TAAAGGGCAG	AACTGCGGCT	ATCGCCGGTG	CCGCGCGAGC	GGTGTACTGA	6850	20
CGACCAGCTG	CGGTAATACC	CTCACATGTT	ACTTGAAGGC	CGCTGCGGCC	6900	
TGTCGAGCTG	CGAAGCTCCA	GGACTGCACG	ATGCTCGTAT	GCGGAGACGA	6950	25
CCTTGTCGTT	ATCTGTGAAA	GCGCGGGGAC	CCAAGAGGAC	GAGGCGAGCC	7000	
TACGGGCCTT	CACGGAGGCT	ATGACTAGAT	ACTCTGCCCC	CCCTGGGGAC	7050	
CCGCCCCAAC	CAGAATACGA	CTTGGAGTTG	ATAACATCAT	GCTCCTCCAA	7100	30
TGTGTGAGTC	GCGCACGATG	CATCTGGCAA	AAGGGTGTAC	TATCTCACCC	7150	
GTGACCCAC	CACCCCCCTT	GCGCGGGCTG	CGTGGGAGAC	AGCTAGACAC	7200	35
ACTCCAGTCA	ATTCCTGGCT	AGGCAACATC	ATCATGTATG	CGCCACCTT	7250	
GTGGGCAAGG	ATGATCCTGA	TGACTCATTT	CTTCTCCATC	CTTCTAGCTC	7300	40
AGGAACAAC	TGAAAAAGCC	CTAGATTGTC	AGATCTACGG	GGCCTGTTAC	7350	
TCCATTGAGC	CACTTGACCT	ACCTCAGATC	ATTCAACGAC	TCCATGGCCT	7400	
TAGCGCATTT	TCACTCCATA	GTTACTCTCC	AGGTGAGATC	AATAGGGTGG	7450	45
CTTCATGCCT	CAGGAACTT	GGGGTACCGC	CCTTGCGAGT	CTGGAGACAT	7500	
CGGGCCAGAA	GTGTCCGCGC	TAGGCTACTG	TCCCAGGGGG	GGAGGGCTGC	7550	50
CACTTGTGGC	AAGTACCTCT	TCAACTGGGC	AGTAAGGACC	AAGCTCAAAC	7600	
TCACTCCAAT	CCCGGCTGCG	TCCCAGTTGG	ATTTATCCAG	CTGGTTCGTT	7650	55
GCTGGTTACA	GCGGGGGAGA	CATATATCAC	AGCCTGTCTC	GTGCCCAGCC	7700	
CCGCTGGTTC	ATGTGGTGCC	TACTCCTACT	TTCTGTAGGG	GTAGGCATCT	7750	60
ATCTACTCCC	CAACCGATGA	ACGGGGAGCT	AAACACTCCA	GGCCAATAGG	7800	
CCATCCTGTT	TTTTTCCCTT	TTTTTTTTTC	TTTTTTTTTT	TTTTTTTTTT	7850	
TTTTTTTTTT	TTTTCTCCTT	TTTTTTTCCT	CTTTTTTTCC	TTTTCTTTCC	7900	65

# DE 199 15 178 A 1

TTTGGTGGCT CCATCTTAGC CCTAGTCACG GCTAGCTGTG AAAGGTCCGT 7950

GAGCCGCTTG ACTGCAGAGA GTGCTGATAC TGGCCTCTCT GCAGATCAAG 8000

5  
T

10

15

20

25

30

35

40

45

50

55

60

65

# DE 199 15 178 A 1

## ANGABEN ZU SEQ ID-NO:4 :

### SEQUENZCHARAKTERISTIKA:

LÄNGE: 7989 Basenpaare

ART: Nukleinsäure

STRANGFORM: Einzelstrang

TOPOLOGIE: linear

ART DES MOLEKÜLS: Subgenom-RNA

HYPOTHETISCH: Nein

ANTI-SENSE: nein

### URSPRÜNGLICHE HERKUNFT:

ORGANISMUS: Hepatitis C Virus

STAMM: Genotyp 1b

ENTWICKLUNGSSTADIUM: —

ZELLTYP: —

### UNMITTELBARE HERKUNFT:

Isolat aus humanen Leberzellen

### MERKMAL:

NAME/SCHLÜSSEL: kodierende Sequenz für Virusproteine

LAGE: —

ERMITTLUNGSMETHODE: . experimentell

## SEQUENZBESCHREIBUNG: SEQ ID-NO: 4

I<sub>377</sub>/NS3-3'/wt: 7989 bp

GCCAGCCCCC GATTGGGGGC GAACTCCAC CATAGATCAC TCCCCTGTGA	50	50
GGAAGTACTG TCTTACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG	100	
TGTCGTGCAG CCTCCAGGAC CCCCCCTCCC GGGAGAGCCA TAGTGGTCTG	150	55
CGGAACCGGT GAGTACACCG GAATTGCCAG GACGACCGGG TCCTTTCTTG	200	
GATCAACCCG CTCAATGCCT GGAGATTGG GCGTGCCCC GCGAGACTGC	250	60
TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG	300	
GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGAC CATGAGCACG	350	
AATCCTAAAC CTCAAAGAAA AACCAAAGGG CGCGCCATGA TTGAACAAGA	400	65

# DE 199 15 178 A 1

	TGGATTGCAC	GCAGGTTCTC	CGGCCGCTTG	GGTGGAGAGG	CTATTCGGCT	450
	ATGACTGGGC	ACAACAGACA	ATCGGCTGCT	CTGATGCCGC	CGTGTTCGG	500
5	CTGTCAGCGC	AGGGGCGCCC	GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	550
	TGCCCTGAAT	GAAGTGCAGG	ACGAGGCAGC	GCGGCTATCG	TGGCTGGCCA	600
10	CGACGGGCGT	TCCTTGCGCA	GCTGTGCTCG	ACGTTGTCAC	TGAAGCGGGA	650
	AGGGACTGGC	TGCTATTGGG	CGAAGTGCCG	GGGCAGGATC	TCCTGTCATC	700
15	TCACCTTGCT	CCTGCCGAGA	AAGTATCCAT	CATGGCTGAT	GCAATGCGGC	750
	GGCTGCATAC	GCTTGATCCG	GCTACCTGCC	CATTCGACCA	CCAAGCGAAA	800
20	CATCGCATCG	AGCGAGCACG	TACTCGGATG	GAAGCCGGTC	TTGTCGATCA	850
	GGATGATCTG	GACGAAGAGC	ATCAGGGGCT	CGCGCCAGCC	GAAGTGTTCG	900
	CCAGGCTCAA	GGCGCGCATG	CCCGACGGCG	AGGATCTCGT	CGTGACCCAT	950
25	GGCGATGCCT	GCTTGCCGAA	TATCATGGTG	GAAAATGGCC	GCTTTTCTGG	1000
	ATTCATCGAC	TGTGGCCGGC	TGGGTGTGGC	GGACCGCTAT	CAGGACATAG	1050
30	CGTTGGCTAC	CCGTGATATT	GCTGAAGAGC	TTGGCGGCGA	ATGGGCTGAC	1100
	CGCTTCCTCG	TGCTTTACGG	TATCGCCGCT	CCCGATTTCG	AGCGCATCGC	1150
35	CTTCTATCGC	CTTCTTGACG	AGTTCTTCTG	AGTTTAAACA	GACCACAACG	1200
	GTTTCCCTCT	AGCGGGATCA	ATTCCGCCCC	TCTCCCTCCC	CCCCCCTAA	1250
40	CGTTACTGGC	CGAAGCCGCT	TGGAATAAGG	CCGGTGTGCG	TTGTCTATA	1300
	TGTTATTTTC	CACCATATTG	CCGTCTTTTG	GCAATGTGAG	GGCCCGGAAA	1350
	CCTGGCCCTG	TCTTCTTGAC	GAGCATTCCT	AGGGGTCTTT	CCCCTCTCGC	1400
45	CAAAGGAATG	CAAGGTCTGT	TGAATGTCGT	GAAGGAAGCA	GTTCTCTCTG	1450
	AAGCTTCTTG	AAGACAAACA	ACGTCTGTAG	CGACCCCTTG	CAGGCAGCGG	1500
50	AACCCCCAC	CTGGCGACAG	GTGCCTCTGC	GGCCAAAAGC	CACGTGTATA	1550
	AGATACACCT	GCAAAGGCGG	CACAACCCCA	GTGCCACGTT	GTGAGTTGGA	1600
55	TAGTTGTGGA	AAGAGTCAAA	TGGCTCTCCT	CAAGCGTATT	CAACAAGGGG	1650
	CTGAAGGATG	CCCAGAAGGT	ACCCCATTTG	ATGGGATCTG	ATCTGGGGCC	1700
60	TCGGTGCACA	TGCTTTACAT	GTGTTTAGTC	GAGGTTAAAA	AACGTCTAGG	1750
	CCCCCGAAC	CACGGGGACG	TGGTTTTCCT	TTGAAAAACA	CGATAATACC	1800
	ATGGCGCCTA	TTACGGCCTA	CTCCCAACAG	ACGCGAGGCC	TACTTGGCTG	1850
65	CATCATCACT	AGCCTCACAG	GCCGGGACAG	GAACCAGGTC	GAGGGGGAGG	1900

# DE 199 15 178 A 1

TCCAAGTGGT	CTCCACCGCA	ACACAATCTT	TCCTGGCGAC	CTGCGTCAAT	1950	
GGCGTGTGTT	GGACTGTCTA	TCATGGTGCC	GGCTCAAAGA	CCCTTGCCGG	2000	
CCCAAAGGGC	CCAATCACCC	AAATGTACAC	CAATGTGGAC	CAGGACCTCG	2050	5
TCGGCTGGCA	AGCGCCCCC	GGGGCGCGTT	CCTTGACACC	ATGCACCTGC	2100	
GGCAGCTCGG	ACCTTTACTT	GGTCACGAGG	CATGCCGATG	TCATTCCGGT	2150	10
GCGCCGGCGG	GGCGACAGCA	GGGGGAGCCT	ACTCTCCCCC	AGGCCCGTCT	2200	
CCTACTTGAA	GGGCTCTTCG	GGCGGTCCAC	TGCTCTGCCC	CTCGGGGCAC	2250	15
GCTGTGGGCA	TCTTTCGGGC	TGCCGTGTGC	ACCCGAGGGG	TTGCGAAGGC	2300	
GGTGGACTTT	GTACCCGTCG	AGTCTATGGA	AACCACTATG	CGGTCCCCGG	2350	20
TCTTCACGGA	CAACTCGTCC	CCTCCGGCCG	TACCGCAGAC	ATTCCAGGTG	2400	
GCCCATCTAC	ACGCCCCCTAC	TGGTAGCGGC	AAGAGCACTA	AGGTGCCGGC	2450	
TGCGTATGCA	GCCCAAGGGT	ATAAGGTGCT	TGTCCTGAAC	CCGTCCGTCG	2500	25
CCGCCACCCT	AGGTTTCGGG	GCGTATATGT	CTAAGGCACA	TGGTATCGAC	2550	
CCTAACATCA	GAACCGGGGT	AAGGACCATC	ACCACGGGTG	CCCCCATCAC	2600	30
GTACTIONACC	TATGGCAAGT	TTCTTGCCGA	CGGTGGTTGC	TCTGGGGGCG	2650	
CCTATGACAT	CATAATATGT	GATGAGTGCC	ACTCAACTGA	CTCGACCACT	2700	35
ATCCTGGGCA	TCGGCACAGT	CCTGGACCAA	GCGGAGACGG	CTGGAGCGCG	2750	
ACTCGTCGTG	CTCGCCACCG	CTACGCCTCC	GGGATCGGTC	ACCGTGCCAC	2800	40
ATCCAAACAT	CGAGGAGGTG	GCTCTGTCCA	GCACTGGAGA	AATCCCCCTT	2850	
TATGGCAAAG	CCATCCCCAT	CGAGACCATC	AAGGGGGGGA	GGCACCTCAT	2900	
TTTCTGCCAT	TCCAAGAAGA	AATGTGATGA	GCTCGCCGCG	AAGCTGTCCG	2950	45
GCCTCGGACT	CAATGCTGTA	GCATATTACC	GGGGCCTTGA	TGTATCCGTC	3000	
ATACCAACTA	GCGGAGACGT	CATTGTGCGA	GCAACGGACG	CTCTAATGAC	3050	50
GGGCTTTACC	GGCGATTTTCG	ACTCAGTGAT	CGACTGCAAT	ACATGTGTCA	3100	
CCCAGACAGT	CGACTTCAGC	CTGGACCCGA	CCTTCACCAT	TGAGACGACG	3150	55
ACCGTGCCAC	AAGACGCGGT	GTCACGCTCG	CAGCGGCGAG	GCAGGACTGG	3200	
TAGGGGCAGG	ATGGGCATTT	ACAGGTTTGT	GACTCCAGGA	GAACGGCCCT	3250	60
CGGGCATGTT	CGATTCTCTG	GTTCTGTGCG	AGTGCTATGA	CGCGGGCTGT	3300	
GCTTGGTACG	AGCTCACGCC	CGCCGAGACC	TCAGTTAGGT	TGCGGGCTTA	3350	
CCTAAACACA	CCAGGGTTGC	CCGTCTGCCA	GGACCATCTG	GAGTTCTGGG	3400	65

# DE 199 15 178 A 1

	AGAGCGTCTT TACAGGCCTC ACCCACATAG ACGCCCATTT CTTGTCCCAG	3450
	ACTAAGCAGG CAGGAGACAA CTTCCCCTAC CTGGTAGCAT ACCAGGCTAC	3500
5	GGTGTGCGCC AGGGCTCAGG CTCCACCTCC ATCGTGGGAC CAAATGTGGA	3550
	AGTGTCTCAT ACGGCTAAAG CCTACGCTGC ACGGGCCAAC GCCCCTGCTG	3600
10	TATAGGCTGG GAGCCGTTCA AAACGAGGTT ACTACCACAC ACCCCATAAC	3650
	CAAATACATC ATGGCATGCA TGTCGGCTGA CCTGGAGGTC GTCACGAGCA	3700
15	CCTGGGTGCT GGTAGGCGGA GTCCTAGCAG CTCTGGCCGC GTATTGCCTG	3750
	ACAACAGGCA GCGTGGTCAT TGTGGGCAGG ATCATCTTGT CCGGAAAGCC	3800
20	GGCCATCATT CCCGACAGGG AAGTCCTTTA CCGGGAGTTC GATGAGATGG	3850
	AAGAGTGCGC CTCACACCTC CCTTACATCG AACAGGGAAT GCAGCTCGCC	3900
	GAACAATTCA AACAGAAGGC AATCGGGTTG CTGCAAACAG CCACCAAGCA	3950
25	AGCGGAGGCT GCTGCTCCCG TGGTGGAAATC CAAGTGGCGG ACCCTCGAAG	4000
	CCTTCTGGGC GAAGCATATG TGAATTTCA TCAGCGGGAT ACAATATTTA	4050
30	GCAGGCTTGT CCACTCTGCC TGGCAACCCC GCGATAGCAT CACTGATGGC	4100
	ATTCACAGCC TCTATCACCA GCCCCTCAC CACCCAACAT ACCCTCCTGT	4150
35	TTAACATCCT GGGGGGATGG GTGGCCGCCC AACTTGCTCC TCCCAGCGCT	4200
	GCTTCTGCTT TCGTAGGCGC CGGCATCGCT GGAGCGGCTG TTGGCAGCAT	4250
40	AGGCCTTGGG AAGGTGCTTG TGGATATTTT GGCAGGTTAT GGAGCAGGGG	4300
	TGGCAGGCGC GCTCGTGGCC TTAAAGGTCA TGAGCGGCGA GATGCCCTCC	4350
	ACCGAGGACC TGGTTAACCT ACTCCCTGCT ATCCTCTCCC CTGGCGCCCT	4400
45	AGTCGTCGGG GTCGTGTGCG CAGCGATACT GCGTCGGCAC GTGGGCCAG	4450
	GGGAGGGGGC TGTGCAGTGG ATGAACCGGC TGATAGCGTT CGCTTCGCGG	4500
50	GGTAACCACG TCTCCCCAC GCACTATGTG CCTGAGAGCG ACGCTGCAGC	4550
	ACGTGTCACT CAGATCCTCT CTAGTCTTAC CATCACTCAG CTGCTGAAGA	4600
55	GGCTTCACCA GTGGATCAAC GAGGACTGCT CCACGCCATG CTCCGGCTCG	4650
	TGGCTAAGAG ATGTTTGGGA TTGGATATGC ACGGTGTTGA CTGATTTCAA	4700
	GACCTGGCTC CAGTCCAAGC TCCTGCCGCG ATTGCCGGGA GTCCCTTCT	4750
60	TCTCATGTCA ACGTGGGTAC AAGGGAGTCT GGCGGGGCGA CGGCATCATG	4800
	CAAACCACCT GCCCATGTGG AGCACAGATC ACCGGACATG TGAAAAACGG	4850
65	TTCCATGAGG ATCGTGGGGC CTAGGACCTG TAGTAACACG TGGCATGGAA	4900



# DE 199 15 178 A 1

CATTCCCCAT	TAACGCGTAC	ACCACGGGCC	CCTGCACGCC	CTCCCCGGCG	4950	
CCAAATTATT	CTAGGGCGCT	GTGGCGGGTG	GCTGCTGAGG	AGTACGTGGA	5000	
GTTTACGCGG	GTGGGGGATT	TCCACTACGT	GACGGGCATG	ACCACTGACA	5050	5
ACGTAAAGTG	CCCGTGTCTAG	GTTCCGGCCC	CCGAATTCTT	CACAGAAGTG	5100	
GATGGGGTGC	GGTTGCACAG	GTACGCTCCA	GCGTGCAAAC	CCCTCCTACG	5150	10
GGAGGAGGTC	ACATTCTCTGG	TCGGGCTCAA	TCAATACCTG	GTTGGGTCAC	5200	
AGCTCCCATG	CGAGCCCGAA	CCGGACGTAG	CAGTGCTCAC	TTCCATGCTC	5250	15
ACCGACCCCT	CCCACATTAC	GGCGGAGACG	GCTAAGCGTA	GGCTGGCCAG	5300	
GGGATCTCCC	CCCTCCTTGG	CCAGCTCATC	AGCTAGCCAG	CTGTCTGCGC	5350	20
CTTCCTTGAA	GGCAACATGC	ACTACCCGTC	ATGACTCCCC	GGACGCTGAC	5400	
CTCATCGAGG	CCAACCTCCT	GTGGCGGCAG	GAGATGGGCG	GGAACATCAC	5450	25
CCGCGTGAG	TCAGAAAATA	AGGTAGTAAT	TTTGGAATCT	TTCGAGCCGC	5500	
TCCAAGCGGA	GGAGGATGAG	AGGGAAGTAT	CCGTCCGGC	GGAGATCCTG	5550	
CGGAGGTCCA	GGAAATTCCC	TCGAGCGATG	CCCATATGGG	CACGCCCAGA	5600	30
TTACAACCCT	CCACTGTTAG	AGTCCTGGAA	GGACCCGGAC	TACGTCCCTC	5650	
CAGTGGTACA	CGGGTGTCCA	TTGCCGCCTG	CCAAGGCCCC	TCCGATACCA	5700	35
CCTCCACGGA	GGAAGAGGAC	GGTTGTCTTG	TCAGAATCTA	CCGTGTCTTC	5750	
TGCCTTGGCG	GAGCTCGCCA	CAAAGACCTT	CGGCAGCTCC	GAATCGTCGG	5800	40
CCGTCGACAG	CGGCACGGCA	ACGGCCTCTC	CTGACCAGCC	CTCCGACGAC	5850	
GGCGACGCGG	GATCCGACGT	TGAGTCGTAC	TCCTCCATGC	CCCCCCTTGA	5900	45
GGGGGAGCCG	GGGGATCCCG	ATCTCAGCGA	CGGGTCTTGG	TCTACCGTAA	5950	
GCGAGGAGGC	TAGTGAGGAC	GTCGTCTGCT	GCTCGATGTC	CTACACATGG	6000	
ACAGGCGCCC	TGATCACGCC	ATGCGCTGCG	GAGGAAACCA	AGCTGCCCAT	6050	50
CAATGCACTG	AGCAACTCTT	TGCTCCGTCA	CCACAACCTG	GTCTATGCTA	6100	
CAACATCTCG	CAGCGCAAGC	CTGCGGCAGA	AGAAGGTCAC	CTTTGACAGA	6150	55
CTGCAGGTCC	TGGACGACCA	CTACCGGGAC	GTGCTCAAGG	AGATGAAGGC	6200	
GAAGGCGTCC	ACAGTTAAGG	CTAAACTTCT	ATCCGTGGAG	GAAGCCTGTA	6250	60
AGCTGACGCC	CCCACATTCG	GCCAGATCTA	AATTTGGCTA	TGGGGCAAAG	6300	
GACGTCCGGA	ACCTATCCAG	CAAGGCCGTT	AACCACATCC	GCTCCGTGTG	6350	
GAAGGACTTG	CTGGAAGACA	CTGAGACACC	AATTGACACC	ACCATCATGG	6400	65

# DE 199 15 178 A 1

CAAAAAATGA GGTTTTCTGC GTCCAACCAG AGAAGGGGGG CCGCAAGCCA 6450  
 GCTCGCCTTA TCGTATTCCC AGATTG GGG GTTCGTGTGT GCGAGAAAAT 6500  
 5 GGCCCTTTAC GATGTGGTCT CCACCCTCCC TCAGGCCGTG ATGGGCTCTT 6550  
 CATACGGATT CCAATACTCT CCTGGACAGC GGGTCGAGTT CCTGGTGAAT 6600  
 10 GCCTGGAAG CGAAGAAATG CCCTATGGGC TTCGCATATG ACACCCGCTG 6650  
 TTTTGA CTCA ACGGTCACTG AGAATGACAT CCGTGTGAG GAGTCAATCT 6700  
 15 ACCAATGTTG TGA CTGGCC CCCGAAGCCA GACAGGCCAT AAGGTCGCTC 6750  
 ACAGAGCGGC TTTACATCGG GGGCCCCCTG ACTAATTCTA AAGGGCAGAA 6800  
 20 CTGCGGCTAT CGCCGGTGCC GCGCGAGCGG TGTACTGACG ACCAGCTGCG 6850  
 GTAATACCCT CACATGTTAC TTGAAGGCCG CTGCGGCCTG TCGAGCTGCG 6900  
 AAGCTCCAGG ACTGCACGAT GCTCGTATGC GGAGACGACC TTGTCGTTAT 6950  
 25 CTGTGAAAGC GCGGGGACCC AAGAGGACGA GGCGAGCCTA CGGGCCTTCA 7000  
 CGGAGGCTAT GACTAGATAC TCTGCCCCC CTGGGGACCC GCCCAAACCA 7050  
 30 GAATACGACT TGGAGTTGAT AACATCATGC TCCTCCAATG TGTCAGTCGC 7100  
 GCACGATGCA TCTGGCAAAA GGGTGTACTA TCTCACCCGT GACCCACCA 7150  
 35 CCCCCCTTGC GCGGGCTGCG TGGGAGACAG CTAGACACAC TCCAGTCAAT 7200  
 TCCTGGCTAG GCAACATCAT CATGTATGCG CCCACCTTGT GGGCAAGGAT 7250  
 40 GATCCTGATG ACTCATTCT TCTCCATCCT TCTAGCTCAG GAACAACTTG 7300  
 AAAAAGCCCT AGATTGTCAG ATCTACGGGG CCTGT TACTC CATTGAGCCA 7350  
 CTGACCTAC CTCAGATCAT TCAACGACTC CATGGCCTTA GCGCATTTTC 7400  
 45 ACTCCATAGT TACTCTCCAG GTGAGATCAA TAGGGTGGCT TCATGCCTCA 7450  
 GGAAACTTGG GGTACCGCCC TTGCGAGTCT GGAGACATCG GGCCAGAAGT 7500  
 50 GTCCGCGCTA GGCTACTGTC CCAGGGGGG AGGGCTGCCA CTTGTGGCAA 7550  
 GTACCTCTTC AACTGGGCAG TAAGGACCAA GCTCAAATC ACTCCAATCC 7600  
 55 CGGCTGCGTC CCAGTTGGAT TTATCCAGCT GGTTCGTTGC TGGTTACAGC 7650  
 GGGGAGACA TATATCACAG CCTGTCTCGT GCGGACCCC GCTGGTTCAT 7700  
 GTGGTGCTA CTCCTACTTT CTGTAGGGGT AGGCATCTAT CTA CTCCCCA 7750  
 60 ACCGATGAAC GGGGAGCTAA ACACTCCAGG CCAATAGGCC ATCCTGTTTT 7800  
 TTTCCTTTT TTTTTTCTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT 7850  
 65 TTCTCCTTTT TTTTCTCTT TTTTTCCTT TTCTTTCCTT TGGTGGCTCC 7900

DE 199 15 178 A 1

ATCTTAGCCC TAGTCACGGC TAGCTGTGAA AGGTCCGTGA GCCGCTTGAC 7950

TGCAGAGAGT GCTGATACTG GCCTCTCTGC AGATCAAGT

5

10

15

20

25

30

35

40

45

50

55

60

65

# DE 199 15 178 A 1

ANGABEN ZU SEQ ID-NO:5 :

5 SEQUENZCHARAKTERISTIKA:  
 LÄNGE: 8649 Basenpaare  
 ART: Nukleinsäure  
 STRANGFORM: Einzelstrang  
 10 TOPOLOGIE: linear

ART DES MOLEKÜLS: Subgenom-RNA

15 HYPOTHETISCH: Nein

ANTI-SENSE: nein

20 URSPRÜNGLICHE HERKUNFT:  
 ORGANISMUS: Hepatitis C Virus  
 STAMM: Genotyp 1b  
 25 ENTWICKLUNGSSTADIUM: —  
 ZELLTYP: —

UNMITTELBARE HERKUNFT:  
 30 Isolat aus humanen Leberzellen

MERKMAL:  
 NAME/SCHLÜSSEL: kodierende Sequenz für Virusproteine  
 35 LAGE: —  
 ERMITTLUNGSMETHODE: . experimentell

40

45 SEQUENZBESCHREIBUNG: SEQ ID-NO: 5

I<sub>389</sub>/NS2-3' /wt: 8649 bp

50	GCCAGCCCCC GATTGGGGGC GACACTCCAC CATAGATCAC TCCCCTGTGA	50
	GGAAGTACTG TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG	100
55	TGTCGTGCAG CCTCCAGGAC CCCCCCTCCC GGGAGAGCCA TAGTGGTCTG	150
	CGGAACCGGT GAGTACACCG GAATTGCCAG GACGACCGGG TCCTTTCTTG	200
	GATCAACCCG CTCAATGCCT GGAGATTTGG GCGTGCCCCC GCGAGACTGC	250
60	TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG	300
	GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGAC CATGAGCACG	350
65	AATCCTAAAC CTCAAAGAAA AACCAAACGT AACACCAACG GGCGCGCCAT	400

# DE 199 15 178 A 1

GATTGAACAA	GATGGATTGC	ACGCAGGTTT	TCCGGCCGCT	TGGGTGGAGA	450	
GGCTATTTCG	CTATGACTGG	GCACAACAGA	CAATCGGCTG	CTCTGATGCC	500	
GCCGTGTTCC	GGCTGTCAGC	GCAGGGGCGC	CCGGTTCTTT	TTGTCAAGAC	550	5
CGACCTGTCC	GGTGCCCTGA	ATGAACTGCA	GGACGAGGCA	GCGCGGCTAT	600	
CGTGGCTGGC	CACGACGGGC	GTTCCCTTGC	CAGCTGTGCT	CGACGTTGTC	650	10
ACTGAAGCGG	GAAGGGACTG	GCTGCTATTG	GGCGAAGTGC	CGGGGCAGGA	700	
TCTCCTGTCA	TCTCACCTTG	CTCCTGCCGA	GAAAGTATCC	ATCATGGCTG	750	15
ATGCAATGCG	GCGGCTGCAT	ACGCTTGATC	CGGCTACCTG	CCCATTCGAC	800	
CACCAAGCGA	AACATCGCAT	CGAGCGAGCA	CGTACTCGGA	TGGAAGCCGG	850	20
TCTTGTCGAT	CAGGATGATC	TGGACGAAGA	GCATCAGGGG	CTCGCGCCAG	900	
CCGAACTGTT	CGCCAGGCTC	AAGGCGCGCA	TGCCCAGACG	CGAGGATCTC	950	25
GTCGTGACCC	ATGGCGATGC	CTGCTTGCCG	AATATCATGG	TGGAAAATGG	1000	
CCGCTTTTCT	GGATTCATCG	ACTGTGGCCG	GCTGGGTGTG	GCGGACCGCT	1050	
ATCAGGACAT	AGCGTTGGCT	ACCCGTGATA	TTGCTGAAGA	GCTTGGCGGC	1100	30
GAATGGGCTG	ACCGCTTCCT	CGTGCTTTAC	GGTATCGCCG	CTCCCGATTG	1150	
GCAGCGCATC	GCCTTCTATC	GCCTTCTTGA	CGAGTTCTTC	TGAGTTTAAA	1200	35
CAGACCACAA	CGGTTTCCCT	CTAGCGGGAT	CAATTCCGCC	CCTCTCCCTC	1250	
CCCCCCCCCT	AACGTTACTG	GCCGAAGCCG	CTTGAATAAA	GGCCGGTGTG	1300	40
CGTTTGCTA	TATGTTATTT	TCCACCATAT	TGCCGTCTTT	TGGCAATGTG	1350	
AGGGCCCGGA	AACCTGGCCC	TGTCTTCTTG	ACGAGCATTC	CTAGGGGTCT	1400	
TTCCCTCTC	GCCAAAGGAA	TGCAAGGTCT	GTTGAATGTC	GTGAAGGAAG	1450	45
CAGTTCCTCT	GGAAGCTTCT	TGAAGACAAA	CAACGTCTGT	AGCGACCCTT	1500	
TGCAGGCAGC	GGAACCCCCC	ACCTGGCGAC	AGGTGCCTCT	GCGGCCAAAA	1550	50
GCCACGTGTA	TAAGATACAC	CTGCAAAGGC	GGCACAACCC	CAGTGCCACG	1600	
TTGTGAGTTG	GATAGTTGTG	GAAAGAGTCA	AATGGCTCTC	CTCAAGCGTA	1650	55
TTCAACAAGG	GGCTGAAGGA	TGCCCAGAAG	GTACCCCAT	GTATGGGATC	1700	
TGATCTGGGG	CCTCGGTGCA	CATGCTTTAC	ATGTGTTTAG	TCGAGGTTAA	1750	
AAAACGTCTA	GGCCCCCGGA	ACCACGGGGA	CGTGGTTTTT	CTTTGAAAAA	1800	60
CACGATAATA	CCATGGACCG	GGAGATGGCA	GCATCGTGCG	GAGGCGCGGT	1850	
TTTCGTAGGT	CTGATACTCT	TGACCTTGTC	ACCGCACTAT	AAGCTGTTCC	1900	65

# DE 199 15 178 A 1

	TCGCTAGGCT	CATATGGTGG	TTACAATATT	TTATCACCAG	GGCCGAGGCA	1950
	CACTTGCAAG	TGTGGATCCC	CCCCCTCAAC	GTTCGGGGGG	GCCGCGATGC	2000
5	CGTCATCCTC	CTCACGTGCG	CGATCCACCC	AGAGCTAATC	TTTACCATCA	2050
	CCAAAATCTT	GCTCGCCATA	CTCGGTCCAC	TCATGGTGCT	CCAGGCTGGT	2100
10	ATAACCAAAG	TGCCGTACTT	CGTGCGCGCA	CACGGGCTCA	TTCGTGCATG	2150
	CATGCTGGTG	CGGAAGGTTG	CTGGGGGTCA	TTATGTCCAA	ATGGCTCTCA	2200
15	TGAAGTTGGC	CGCACTGACA	GGTACGTACG	TTTATGACCA	TCTCACCCCA	2250
	CTGCGGGACT	GGGCCCACGC	GGGCCTACGA	GACCTTGCGG	TGGCAGTTGA	2300
20	GCCCGTCGTC	TTCTCTGATA	TGGAGACCAA	GGTTATCACC	TGGGGGGCAG	2350
	ACACCGCGGC	GTGTGGGGAC	ATCATCTTGG	GCCTGCCCCG	CTCCGCCCGC	2400
	AGGGGGAGGG	AGATACATCT	GGGACCGGCA	GACAGCCTTG	AAGGGCAGGG	2450
25	GTGGCGACTC	CTCGCGCCTA	TTACGGCCTA	CTCCCAACAG	ACGCGAGGCC	2500
	TACTTGGCTG	CATCATCACT	AGCCTCACAG	GCCGGGACAG	GAACCAGGTC	2550
30	GAGGGGGAGG	TCCAAGTGGT	CTCCACCGCA	ACACAATCTT	TCCTGGCGAC	2600
	CTGCGTCAAT	GGCGTGTGTT	GGACTGTCTA	TCATGGTGCC	GGCTCAAAGA	2650
35	CCCTTGCCGG	CCCAAAGGGC	CCAATCACCC	AAATGTACAC	CAATGTGGAC	2700
	CAGGACCTCG	TCGGCTGGCA	AGCGCCCCCC	GGGGCGCGTT	CCTTGACACC	2750
40	ATGCACCTGC	GGCAGCTCGG	ACCTTTACTT	GGTCACGAGG	CATGCCGATG	2800
	TCATTCCGGT	GCGCCGGCGG	GGCGACAGCA	GGGGGAGCCT	ACTCTCCCCC	2850
	AGGCCCCTCT	CCTACTTGAA	GGGCTCTTCG	GGCGGTCCAC	TGCTCTGCCC	2900
45	CTCGGGGCAC	GCTGTGGGCA	TCTTTCGGGC	TGCCGTGTGC	ACCCGAGGGG	2950
	TTGCGAAGGC	GGTGGACTTT	GTACCCGTCG	AGTCTATGGA	AACCACTATG	3000
50	CGGTCCCCGG	TCTTCACGGA	CAACTCGTCC	CCTCCGGCCG	TACCGCAGAC	3050
	ATTCCAGGTG	GCCCATCTAC	ACGCCCCCTAC	TGGTAGCGGC	AAGAGCACTA	3100
55	AGGTGCCGGC	TGCGTATGCA	GCCCAAGGGT	ATAAGGTGCT	TGTCCTGAAC	3150
	CCGTCCGTCG	CCGCCACCCT	AGGTTTCGGG	GCGTATATGT	CTAAGGCACA	3200
	TGGTATCGAC	CCTAACATCA	GAACCGGGGT	AAGGACCATC	ACCACGGGTG	3250
60	CCCCCATCAC	GTA CTCCACC	TATGGCAAGT	TTCTTGCCGA	CGGTGGTTGC	3300
	TCTGGGGGCG	CCTATGACAT	CATAATATGT	GATGAGTGCC	ACTCAACTGA	3350
65	CTCGACCACT	ATCCTGGGCA	TCGGCACAGT	CCTGGACCAA	GCGGAGACGG	3400

# DE 199 15 178 A 1

CTGGAGCGCG	ACTCGTCGTG	CTCGCCACCG	CTACGCCTCC	GGGATCGGTC	3450	
ACCGTGCCAC	ATCCAAACAT	CGAGGAGGTG	GCTCTGTCCA	GCACTGGAGA	3500	
AATCCCCTTT	TATGGCAAAG	CCATCCCCAT	CGAGACCATC	AAGGGGGGGA	3550	5
GGCACCTCAT	TTTCTGCCAT	TCCAAGAAGA	AATGTGATGA	GCTCGCCGCG	3600	
AAGCTGTCCG	GCCTCGGACT	CAATGCTGTA	GCATATTACC	GGGGCCTTGA	3650	10
TGTATCCGTC	ATACCAACTA	GCGGAGACGT	CATTGTCGTA	GCAACGGACG	3700	
CTCTAATGAC	GGGCTTTACC	GGCGATTTTCG	ACTCAGTGAT	GCACTGCAAT	3750	15
ACATGTGTCA	CCCAGACAGT	CGACTTCAGC	CTGGACCCGA	CCTTCACCAT	3800	
TGAGACGACG	ACCGTGCCAC	AAGACGCGGT	GTCACGCTCG	CAGCGGCGAG	3850	20
GCAGGACTGG	TAGGGGCAGG	ATGGGCATTT	ACAGGTTTGT	GACTCCAGGA	3900	
GAACGGCCCT	CGGGCATGTT	CGATTCCCTCG	GTTCTGTGCG	AGTGCTATGA	3950	
CGCGGGCTGT	GCTTGGTACG	AGCTCACGCC	CGCCGAGACC	TCAGTTAGGT	4000	25
TGCGGGCTTA	CCTAAACACA	CCAGGGTTGC	CCGTCTGCCA	GGACCATCTG	4050	
GAGTTCTGGG	AGAGCGTCTT	TACAGGCCTC	ACCCACATAG	ACGCCCATTT	4100	30
CTTGTCCCAG	ACTAAGCAGG	CAGGAGACAA	CTTCCCCTAC	CTGGTAGCAT	4150	
ACCAGGCTAC	GGTGTGCGCC	AGGGCTCAGG	CTCCACCTCC	ATCGTGGGAC	4200	35
CAAATGTGGA	AGTGTCTCAT	ACGGCTAAAG	CCTACGCTGC	ACGGGCCAAC	4250	
GCCCCTGCTG	TATAGGCTGG	GAGCCGTTCA	AAACGAGGTT	ACTACCACAC	4300	40
ACCCCATAAAC	CAAATACATC	ATGGCATGCA	TGTCGGCTGA	CCTGGAGGTC	4350	
GTCACGAGCA	CCTGGGTGCT	GGTAGGCGGA	GTCCTAGCAG	CTCTGGCCGC	4400	
GTATTGCCTG	ACAACAGGCA	GCGTGGTCAT	TGTGGGCAGG	ATCATCTTGT	4450	45
CCGGAAAGCC	GGCCATCATT	CCCGACAGGG	AAGTCCTTTA	CCGGGAGTTC	4500	
GATGAGATGG	AAGAGTGCGC	CTCACACCTC	CCTTACATCG	AACAGGGAAT	4550	50
GCAGCTCGCC	GAACAATTCA	AACAGAAGGC	AATCGGGTTG	CTGCAAACAG	4600	
CCACCAAGCA	AGCGGAGGCT	GCTGCTCCCG	TGGTGGAATC	CAAGTGGCGG	4650	55
ACCCCTGAAG	CCTTCTGGGC	GAAGCATATG	TGGAATTTCA	TCAGCGGGAT	4700	
ACAATATTTA	GCAGGCTTGT	CCACTCTGCC	TGGCAACCCC	GCGATAGCAT	4750	60
CACTGATGGC	ATTCACAGCC	TCTATACCA	GCCCCTCAC	CACCCAACAT	4800	
ACCCTCCTGT	TTAACATCCT	GGGGGGATGG	GTGGCCGCC	AACTTGCTCC	4850	
TCCCAGCGCT	GCTTCTGCTT	TCGTAGGCGC	CGGCATCGCT	GGAGCGGCTG	4900	65

# DE 199 15 178 A 1

	TTGGCAGCAT AGGCCTTGGG AAGGTGCTTG TGGATATTTT GGCAGGTTAT	4950
	GGAGCAGGGG TGGCAGGCGC GCTCGTGGCC TTTAAGGTCA TGAGCGGCGA	5000
5	GATGCCCTCC ACCGAGGACC TGGTTAACCT ACTCCCTGCT ATCCTCTCCC	5050
	CTGGCGCCCT AGTCGTCGGG GTCGTGTGCG CAGCGATACT GCGTCGGCAC	5100
10	GTGGGCCCAG GGGAGGGGGC TGTGCAGTGG ATGAACCGGC TGATAGCGTT	5150
	CGCTTCGCGG GGTAACCACG TCTCCCCAC GCACTATGTG CCTGAGAGCG	5200
15	ACGCTGCAGC ACGTGTCACT CAGATCCTCT CTAGTCTTAC CATCACTCAG	5250
	CTGCTGAAGA GGCTTCACCA GTGGATCAAC GAGGACTGCT CCACGCCATG	5300
20	CTCCGGCTCG TGGCTAAGAG ATGTTTGGGA TTGGATATGC ACGGTGTTGA	5350
	CTGATTTCAA GACCTGGCTC CAGTCCAAGC TCCTGCCGCG ATTGCCGGA	5400
	GTCCCCTTCT TCTCATGTCA ACGTGGGTAC AAGGGAGTCT GCGGGGCGA	5450
25	CGGCATCATG CAAACCACCT GCCCATGTGG AGCACAGATC ACCGACATG	5500
	TGAAAAACGG TTCCATGAGG ATCGTGGGGC CTAGGACCTG TAGTAACACG	5550
30	TGGCATGGAA CATTCCCCAT TAACGCGTAC ACCACGGGCC CCTGCACGCC	5600
	CTCCCCGGCG CCAAATTATT CTAGGGCGCT GTGGCGGGTG GCTGCTGAGG	5650
35	AGTACGTGGA GGTACGCGG GTGGGGGATT TCCACTACGT GACGGGCATG	5700
	ACCACTGACA ACGTAAAGTG CCCGTGTCAG GTTCCGGCCC CCGAATTCTT	5750
40	CACAGAAGTG GATGGGGTGC GGTGTCACAG GTACGCTCCA GCGTGCAAAC	5800
	CCCTCCTACG GGAGGAGGTC ACATTCCCTGG TCGGGCTCAA TCAATACCTG	5850
	GTGGGGTCAC AGCTCCCATG CGAGCCCGAA CCGGACGTAG CAGTGCTCAC	5900
45	TTCCATGCTC ACCGACCCCT CCCACATTAC GCGGAGACG GCTAAGCGTA	5950
	GGCTGGCCAG GGGATCTCCC CCCTCCTTGG CCAGCTCATC AGCTAGCCAG	6000
50	CTGTCTGCGC CTTCTTGAA GGCAACATGC ACTACCCGTC ATGACTCCCC	6050
	GGACGCTGAC CTCATCGAGG CCAACCTCCT GTGGCGGCAG GAGATGGGCG	6100
55	GGAACATCAC CCGCGTGGAG TCAGAAAATA AGGTAGTAAT TTTGGACTCT	6150
	TTGAGCCGC TCCAAGCGGA GGAGGATGAG AGGGAAGTAT CCGTTCCGGC	6200
	GGAGATCCTG CGGAGGTCCA GGAAATTCCC TCGAGCGATG CCCATATGGG	6250
60	CACGCCCCGA TTACAACCCT CCACTGTTAG AGTCCTGGAA GGACCCGGAC	6300
	TACGTCCCTC CAGTGGTACA CGGGTGTCCA TTGCCGCTG CCAAGGCCCC	6350
65	TCCGATACCA CCTCCACGGA GGAAGAGGAC GGTGTCTCTG TCAGAATCTA	6400



# DE 199 15 178 A 1

CCGTGTCTTC	TGCCTTGGCG	GAGCTCGCCA	CAAAGACCTT	CGGCAGCTCC	6450	
GAATCGTCGG	CCGTCGACAG	CGGCACGGCA	ACGGCCTCTC	CTGACCAGCC	6500	
CTCCGACGAC	GGCGACGCGG	GATCCGACGT	TGAGTCGTAC	TCCTCCATGC	6550	5
CCCCCCTTGA	GGGGGAGCCG	GGGGATCCCG	ATCTCAGCGA	CGGGTCTTGG	6600	
TCTACCGTAA	GCGAGGAGGC	TAGTGAGGAC	GTCGTCTGCT	GCTCGATGTC	6650	10
CTACACATGG	ACAGGCGCCC	TGATCACGCC	ATGCGCTGCG	GAGGAAACCA	6700	
AGCTGCCCCAT	CAATGCACTG	AGCAACTCTT	TGCTCCGTCA	CCACAACCTG	6750	15
GTCTATGCTA	CAACATCTCG	CAGCGCAAGC	CTGCGGCAGA	AGAAGGTCAC	6800	
CTTTGACAGA	CTGCAGGTCC	TGGACGACCA	CTACCGGGAC	GTGCTCAAGG	6850	20
AGATGAAGGC	GAAGGCGTCC	ACAGTTAAGG	CTAAACTTCT	ATCCGTGGAG	6900	
GAAGCCTGTA	AGCTGACGCC	CCCACATTCG	GCCAGATCTA	AATTTGGCTA	6950	25
TGGGGCAAAG	GACGTCCGGA	ACCTATCCAG	CAAGGCCGTT	AACCACATCC	7000	
GCTCCGTGTG	GAAGGACTTG	CTGGAAGACA	CTGAGACACC	AATTGACACC	7050	
ACCATCATGG	CAAAAAATGA	GGTTTTCTGC	GTCCAACCAG	AGAAGGGGGG	7100	30
CCGCAAGCCA	GCTCGCCTTA	TCGTATTCCC	AGATTTGGGG	GTTCGTGTGT	7150	
GCGAGAAAAT	GGCCCTTTAC	GATGTGGTCT	CCÀCCCTCCC	TCAGGCCGTG	7200	35
ATGGGCTCTT	CATACGGATT	CCAATACTCT	CCTGGACAGC	GGGTCGAGTT	7250	
CCTGGTGAAT	GCCTGGAAAG	CGAAGAAATG	CCCTATGGGC	TTCGCATATG	7300	40
ACACCCGCTG	TTTTGACTCA	ACGGTCACTG	AGAATGACAT	CCGTGTTGAG	7350	
GAGTCAATCT	ACCAATGTTG	TGACTTGGCC	CCCGAAGCCA	GACAGGCCAT	7400	
AAGGTCGCTC	ACAGAGCGGC	TTTACATCGG	GGGCCCCCTG	ACTAATTCTA	7450	45
AAGGGCAGAA	CTGCGGCTAT	CGCCGGTGCC	GCGCGAGCGG	TGTA CTGACG	7500	
ACCAGCTGCG	GTAATACCCT	CACATGTTAC	TTGAAGGCCG	CTGCGGCCTG	7550	50
TCGAGCTGCG	AAGCTCCAGG	ACTGCACGAT	GCTCGTATGC	GGAGACGACC	7600	
TTGTCTGTTAT	CTGTGAAAGC	GCGGGGACCC	AAGAGGACGA	GGCGAGCCTA	7650	55
CGGGCCTTCA	CGGAGGCTAT	GACTAGATAC	TCTGCCCCCC	CTGGGGACCC	7700	
GCCCCAACCA	GAATACGACT	TGGAGTTGAT	AACATCATGC	TCCTCCAATG	7750	
TGTCAGTCGC	GCACGATGCA	TCTGGCAAAA	GGGTGTACTA	TCTCACCCGT	7800	60
GACCCACCA	CCCCCCTTGC	GCGGGCTGCG	TGGGAGACAG	CTAGACACAC	7850	
TCCAGTCAAT	TCCTGGCTAG	GCAACATCAT	CATGTATGCG	CCCACCTTGT	7900	65

GGGCAAGGAT GATCCTGATG ACTCATTTCT TCTCCATCCT TCTAGCTCAG 7950  
 GAACAACCTTG AAAAAGCCCT AGATTGTCAG ATCTACGGGG CCTGTTACTC 8000  
 5 CATTGAGCCA CTTGACCTAC CTCAGATCAT TCAACGACTC CATGGCCTTA 8050  
 GCGCATTTTC ACTCCATAGT TACTCTCCAG GTGAGATCAA TAGGGTGGCT 8100  
 10 TCATGCCTCA GGAAACTTGG GGTACCGCCC TTGCGAGTCT GGAGACATCG 8150  
 GGCCAGAAGT GTCCGCGCTA GGCTACTGTC CCAGGGGGGG AGGGCTGCCA 8200  
 15 CTTGTGGCAA GTACCTCTTC AACTGGGCAG TAAGGACCAA GCTCAAATC 8250  
 ACTCCAATCC CGGCTGCGTC CCAGTTGGAT TTATCCAGCT GGTTCGTTGC 8300  
 TGGTTACAGC GGGGGAGACA TATATCACAG CCTGTCTCGT GCCCAGCCCC 8350  
 20 GCTGGTTCAT GTGGTGCCTA CTCCTACTTT CTGTAGGGGT AGGCATCTAT 8400  
 CTAATCCCCA ACCGATGAAC GGGGAGCTAA AACTCCAGG CCAATAGGCC 8450  
 25 ATCTGTGTTT TTTCCCTTTT TTTTCTCTT TTTTCTCTT TTTTCTCTT 8500  
 TTTTCTCTT TTTTCTCTT TTTTCTCTT TTTTCTCTT TTTTCTCTT 8550  
 30 TGGTGGCTCC ATCTTAGCCC TAGTCACGGC TAGCTGTGAA AGGTCCGTGA 8600  
 GCGGCTTGAC TGCAGAGAGT GCTGATACTG GCCTCTCTGC AGATCAAGT

# 35 Patentansprüche

1. Hepatitis C Virus (HCV) Zellkultursystem, das im wesentlichen eukaryontische Zellen umfaßt, die eingeschleustes HCV-spezifisches Genmaterial enthalten, **dadurch gekennzeichnet**, daß die eukaryontischen Zellen humane Hepatomazellen sind und daß das eingeschleuste HCV-spezifische Genmaterial ein HCV-RNA-Konstrukt ist, das die HCV-spezifischen RNA-Abschnitte 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3'NTR und zusätzlich ein selektierbares Markergen (Selektionsgen) umfaßt.
2. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet, daß die Hepatomazellen von einer handelsüblichen Hepatomazelllinie abstammen.
3. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet, daß die Hepatomazellen aus einer Hepatomaprimärzellkultur gewonnen sind.
4. HCV-RNA-Konstrukt, dadurch gekennzeichnet, daß es die HCV-spezifischen RNA-Abschnitte 5'NTR, NS3, NS4A, NS4B, NS5A, NS5B und 3'NTR umfaßt.
5. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet, daß es die in Sequenzprotokoll SEQ ID NO: 1 dargestellte Nukleotidsequenz umfaßt.
6. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet, daß es die in Sequenzprotokoll SEQ ID NO: 2 dargestellte Nukleotidsequenz umfaßt.
7. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet, daß es die in Sequenzprotokoll SEQ ID NO: 3 dargestellte Nukleotidsequenz umfaßt.
8. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet, daß es die in Sequenzprotokoll SEQ ID NO: 4 dargestellte Nukleotidsequenz umfaßt.
9. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet, daß es die in Sequenzprotokoll SEQ ID NO: 5 dargestellte Nukleotidsequenz umfaßt.
10. HCV-RNA-Konstrukt nach Anspruch 4, dadurch gekennzeichnet, daß die 3'NTR eine Nukleotidsequenz aufweist, die aus der Gruppe der nachfolgend aufgelisteten Nukleotidsequenzen (a) bis (i) ausgewählt ist:

- (a) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT  
TTAGCTTTTTTTTTTTCTTTTTTTTTTGAGAGAGAGAGTCTCACTCTG  
TTGCCCAGACTGGAGT 5
- (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
CTCACTCTGT TGCCCAGACT GGAGC 10
- (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTAATCTTT TTTTTTTTTCT TTTTTTTTGA GAGAGAGAGT  
CTCACTCTGT TGCCCAGACT GCAGC 15
- (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG  
AGTCTCACTC TGTTGCCCAG ACTGGAGT 20
- (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG  
AGTCTCACTC TGTTGCCCAG ACTGGAGT 25
- (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTT TCTTTTTTTT TTGAGAGAGA  
GAGTCTCACT CTGTTGCCCA GACTGGAGT 30
- (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTAGTCT TTTTTTTTTT CTTTTTTTTT GAGAGAGAGA  
GTCTCACTCT GTTGCCCAGA CTGGAGT 35
- (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTTAAT CTTTTTTTTT TTTTCTTTT TTTTGAGAGA  
GAGAGTCTCA CTCTGTTGCC CAGACTGGAG T 40
- (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTT  
TTTTTTAATC TTTTTTTTTT TTTTCTTTT TTTTGAGAG  
AGAGAGTCTC ACTCTGTTGC CCAGACTGGA GT 45

11. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet, daß das selektierbare Markergen ein Resistenzgen und insbesondere eine Antibiotikaresistenzgen ist.
12. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 10, dadurch gekennzeichnet, daß das selektierbare Markergen ein Neomycinphosphotransferasegen ist.
13. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 12, dadurch gekennzeichnet, daß das selektierbare Markergen strangabwärts der 5'NTR in die HCV-RNA integriert ist.
14. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 13, dadurch gekennzeichnet, daß das selektierbare Markergen über ein Ribozym bzw. eine Erkennungsstelle für ein Ribozym mit der HCV-RNA verbunden ist.
15. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 14, dadurch gekennzeichnet, daß es ein integriertes Reportergen aufweist.
16. HCV-RNA-Konstrukt nach Anspruch 15, dadurch gekennzeichnet, daß das Reportergen ein Gen aus der Gruppe der Luziferasegene, dem CAT-Gen (Chloramphenicol-Acetyl-Transferase-Gen), dem lacZ-Gen (beta-Galaktosidasegen), der GFP-Gene (green-fluorescence-protein-Gene), dem GUS-Gen (Glukuronidasegen) und dem SEAP-Gen (Sezernierte-Alkalische-Phosphatase-Gen) ist.
17. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 15, dadurch gekennzeichnet, daß deren Replikation die Expression eines (zellulären) Surrogatmarkergens beeinflußt.
18. HCV-RNA-Konstrukt nach einem der Ansprüche 15 bis 17, dadurch gekennzeichnet, daß das Resistenzgen derart in das offene Leseraster der HCV-RNA einidioniert ist, daß es erst nach einer proteolytischen Prozessierung in eine aktive Form überführbar ist.
19. HCV-RNA-Konstrukt nach einem der Ansprüche 15 bis 18, dadurch gekennzeichnet, daß das Reportergen und das selektierbare Markergen derart räumlich in dem Konstrukt angeordnet sind, daß sie gemeinsam ein Fusionspro-

tein exprimieren.

20. Zellkultursystem nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß das HCV-RNA-Konstrukt ein Konstrukt gemäß wenigstens einem der Ansprüche 4 bis 19 ist.

21. Zellkultursystem nach Anspruch 1, dadurch gekennzeichnet, daß die das HCV-RNA-Konstrukt enthaltenden Zellen bei der DSMZ, Braunschweig, BRD, unter der Hinterlegungsnummer DSM ACC2394 (Laborbezeichnung HuBI 9-13) hinterlegt sind.

22. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 20 bis 21 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 19 zur Herstellung und/oder Evaluierung und/oder Testung von Therapeutika und/oder Diagnostika zur Behandlung von insbesondere HCV-Infektionen.

23. Verwendung eines Zellkultursystems nach einem der Ansprüche 1 bis 3 oder 20 bis 21 und/oder eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 19 zur Herstellung eines Impfstoffes gegen HCV-Infektionen.

24. Verwendung eines HCV-RNA-Konstrukts nach einem der Ansprüche 4 bis 19 zur Herstellung einer leberzell-spezifischen Genfähe für die Gentherapie.

25. HCV-RNA-Konstrukt nach einem der Ansprüche 4 bis 19, dadurch gekennzeichnet, daß es ein integriertes Fremdgen aufweist und dazu geeignet ist, dieses Fremdgen in eine Zielzelle einzuschleusen, die zur Expression dieses Fremdgens geeignet ist.

---

Hierzu 6 Seite(n) Zeichnungen

---

20

25

30

35

40

45

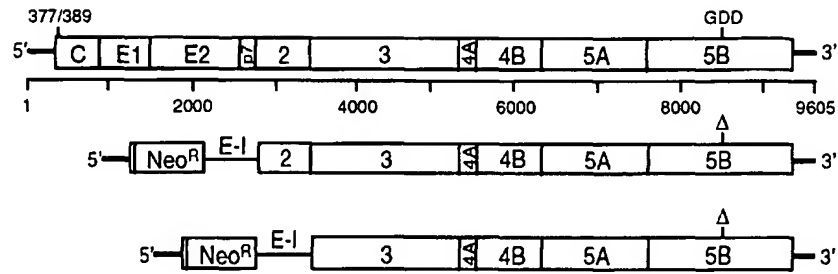
50

55

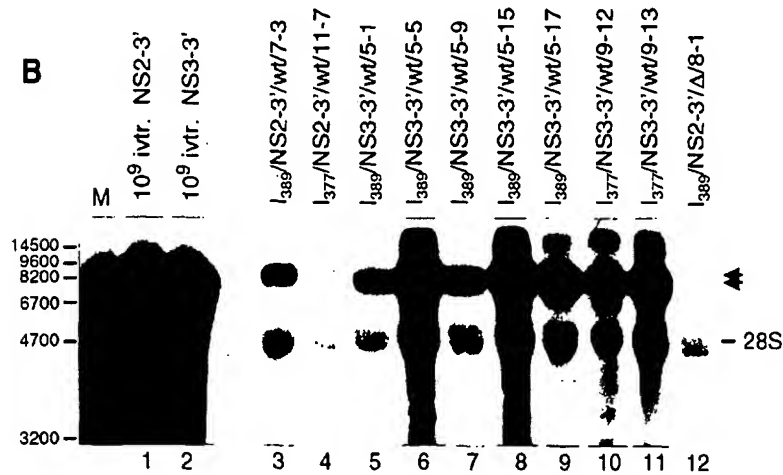
60

65

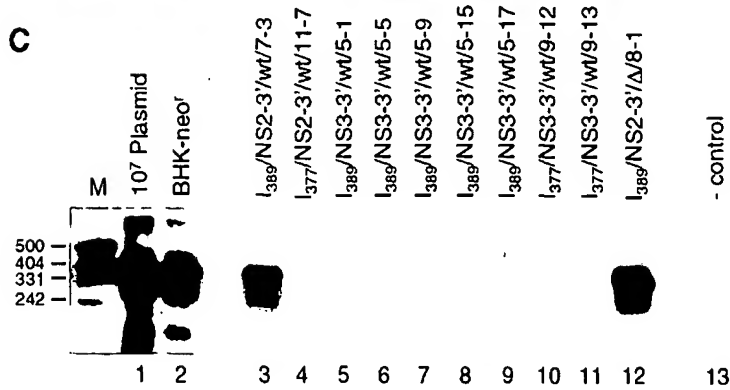
**A**



**B**



**C**



**Fig. 1**

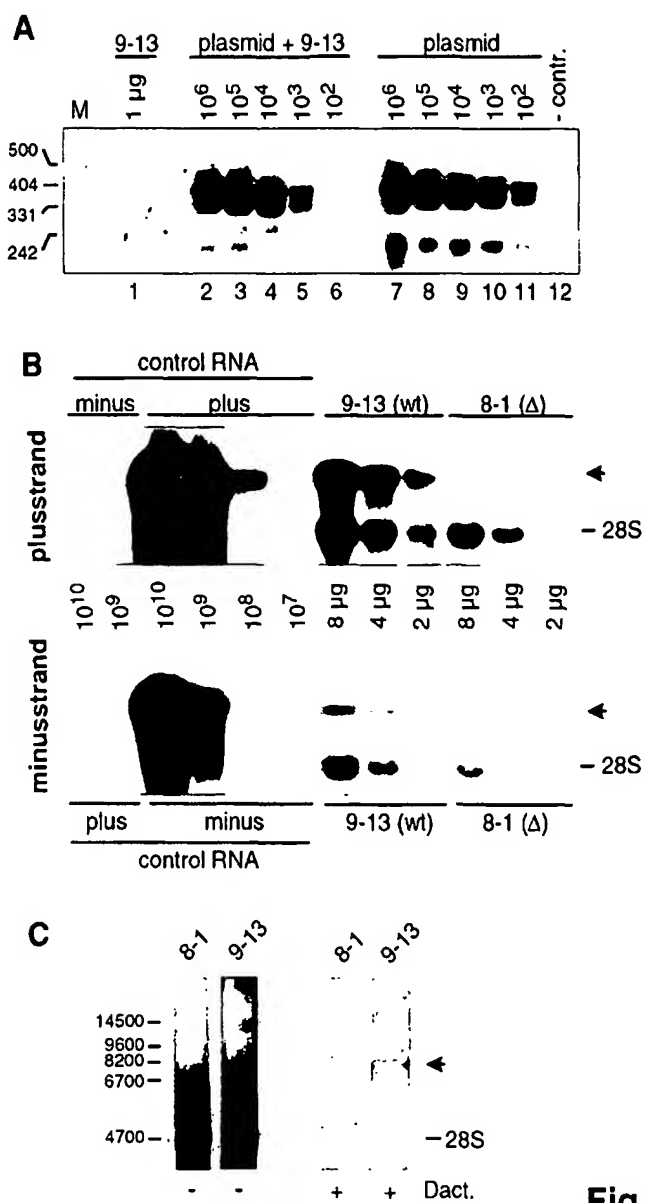


Fig. 2

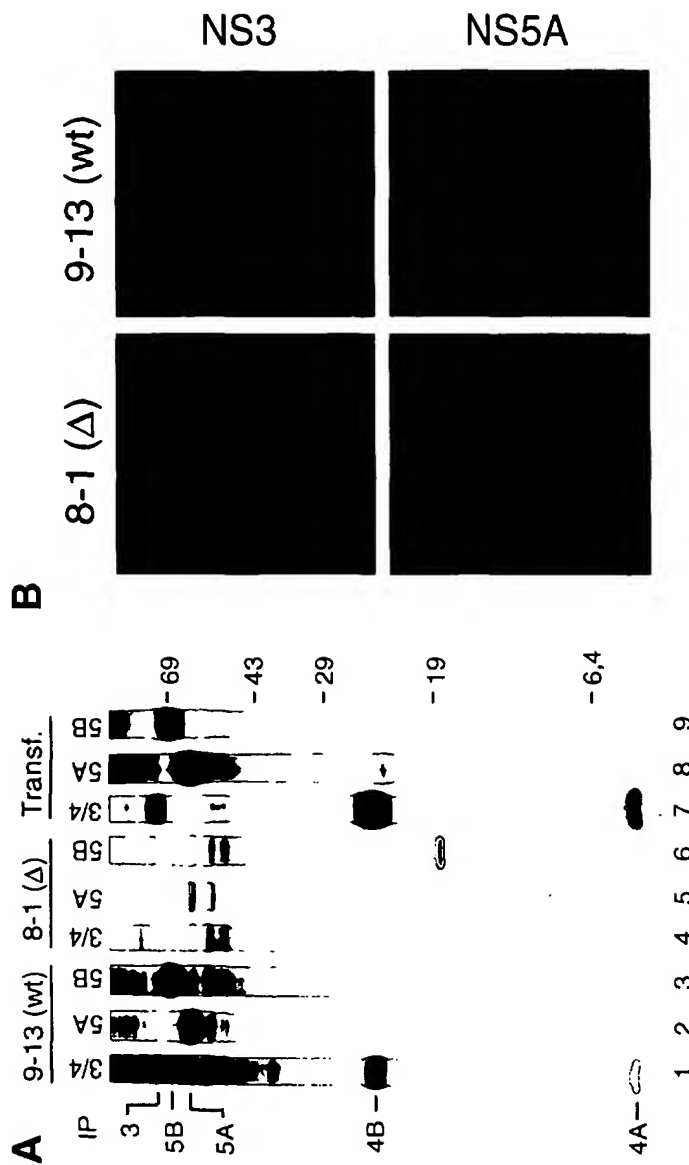
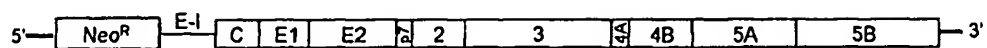
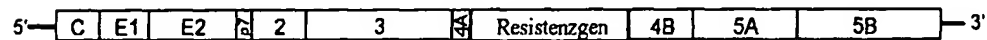
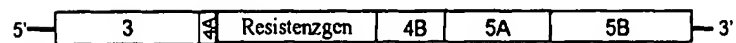


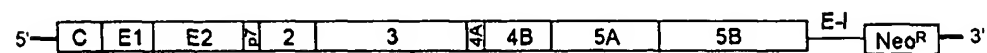
Fig. 3



**Fig. 4**



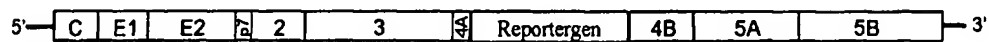
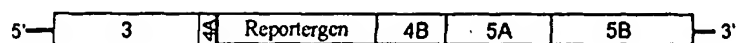
A



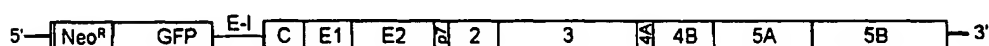
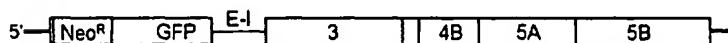
B

**Fig. 5**

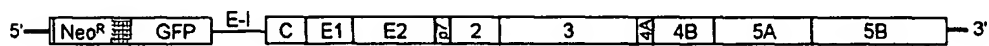
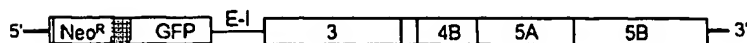




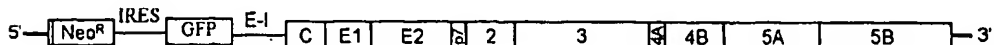
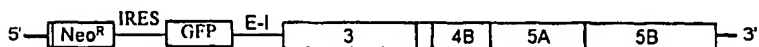
A



B

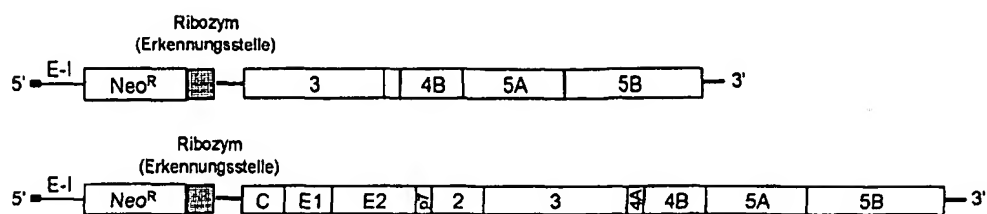


C

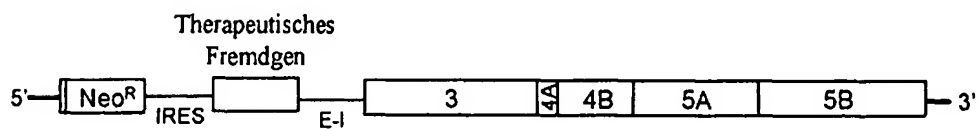


D

**Fig. 6**



**Fig. 7**



**Fig. 8**